View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

SITE-SATURATION MUTAGENESIS OF CRUCIAL RESIDUES IN NAD / NADP SPECIFICITY OF FORMATE DEHYDROGENASE FROM *Candida methylica*

M.Sc. Thesis by
Gülşah PAR

Department : **Advanced Technologies**

Programme : **Molecular Biology – Genetics and Biotechnology**

JUNE 2009

İSTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

SITE-SATURATION MUTAGENESIS OF CRUCIAL RESIDUES IN NAD / NADP SPECIFICITY OF FORMATE DEHYDROGENASE FROM *Candida methylica*

M.Sc. Thesis by
Gülşah PAR
(521061212)

Date of submission : 04 May 2009
Date of defence examination: 05 June 2009

Supervisor (Chairman) : Assist. Prof. Dr. Nevin GÜL KARAGÜLER (ITU)
Members of the Examining Committee : Assist. Prof. Dr. Gizem DİNLER (ITU)
Assist. Prof. Dr. Nagehan ERSOY (HU)

JUNE 2009

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

Candida methylica FORMAT DEHİDROGENAZ'IN NAD / NADP SPESİFİTESİ İÇİN KRİTİK BÖLGELERİN BÖLGE SATURASYON MUTAGENEZİ

YÜKSEK LİSANS TEZİ

Gülşah PAR
(521061212)

Tezin Enstitüye Verildiği Tarih : 04 Mayıs 2009
Tezin Savunulduğu Tarih : 05 Haziran 2009

Tez Danışmanı : Yrd.Doç. Dr. Nevin GÜL KARAGÜLER(İTÜ)
Diğer Jüri Üyeleri : Yrd.Doç. Dr. Gizem DİNLER (İTÜ)
Yrd.Doç. Dr. Nagehan ERSOY (HÜ)

HAZİRAN 2009

FOREWORD

I would like to thank and express my sincere gratitude to my thesis supervisor Assist. Prof. Nevin Gül Karagüler for her guidance and contribution throughout my graduation study.

I would like to acknowledge and thank to my lab partners for their collaboration in ITU PROTEIN ENGINEERING LAB., especially PhD students Emel Ordu and Emrah Yelboga.

I acknowledge the financial support from the Scientific and Technological Research Council of Turkey (TÜBİTAK) through the project with grant number 107T684.

Finally, I wish to express my love and gratitude to my beloved family; for their endless support and love. They taught me the values that have helped me to become the individual that I am today.

June 2009

Gülşah PAR

Biologist

TABLE OF CONTENTS

Page

ABBREVIATIONS

***cb*FDH**	**:** *Candida boidinii* Formate Dehydrogenase
***cm*FDH**	**:** *Candida methylica* Formate Dehydrogenase
EDTA	**:** Ethylenediaminetetraaceticacid
GC	**:** Gas Chromatography
HPLC	**:** High Performance Liquid Chromatography
IPTG	**:** Isopropyl-Beta-D-Thiogalactopyranoside
LB	**:** Luria-Bertani Broth (Lysogeny Broth)
MS	**:** Mass Spectrometry
NAD	**:** Nicotinamide Adenine Dinucleotide
NADP	**:** Nicotinamide Adenine Dinucleotide Phosphate
NBT	**:** Nitroblue Tetrozolium
OD	**:** Optic Density
PMS	**:** Phenazine Methosulfate
***ps*FDH**	**:** *Pseudomonas sp101*Formate Dehydrogenase
RT	**:** Room Temperature
***sc*FDH**	**:** *Saccharomyces cerevisiae* Formate Dehydrogenase
SeSaM	**:** Sequence Saturation Mutagenesis
SOC	**:** Super Optimal Broth with Catabolite Repression
SSM	**:** Site Saturation Mutagenesis
TBE	**:** Tris/Borate/EDTA Buffer

LIST OF TABLES

Page

LIST OF FIGURES

Page

SITE-SATURATION MUTAGENESIS OF CRUCIAL RESIDUES IN NAD^+ / $NADP^+$ SPECIFICITY OF FORMATE DEHYDROGENASE FROM *Candida methylica*

SUMMARY

NAD^+-dependent formate dehydrogenase (EC 1.2.1.2, FDH) plays an important role in the final step of methanol oxidation pathway in methylotrophs, which catalyzes the oxidation of formate anion into carbondioxide, coupled with reduction of NAD^+ to NADH. However FDH is widely used in regeneration of NADH, it could not been used in NADPH regeneration for the reason that all FDHs which are already found in nature are highly specific to only NAD^+ coenzyme. On the other hand NADPH regeneration via this enzyme using $NADP^+$ as a coenzyme is very desirable because of the advantages of FDH, such as its availability and low cost, a favourable thermodynamic equilibrium and the inertness of CO_2. In order to overcome limited coenzyme specifity of FDH, many attempts using a rational design approach have been made recently. However all these mutants bind $NADP^+$ very weakly and still show activity with NAD^+.

In this project, we used site saturation mutagenesis, which is a technique using a directed evolution approach to redesign of proteins to improve the K_M of *Candida methylica* FDH for $NADP^+$. Firstly, in the coenzyme binding domain, the amino acid residues which are responsible for the coenzyme specificity were determined by using Insight II (Accelrys) program on a homology model of *cm*FDH based on *Pseudomonas. sp*.101 and *Candida boidinii* FDH crystal structure. After the application of site saturation mutagenesis with degenerate primers to the determined residues D195, Y196, Q197, mutant libraries were constructed. The efficient mutants have been screened by using colorimetric screening assay. 27 candidates have shown activity towards $NADP^+$ from nearly 400 screened colonies. These results reveal that, these residues are important in controlling the cofactor specificity of formate dehydrogenases, and promising for developing $NADP^+$-dependent *cm*FDH enzyme.
In order to determine the most active and effective candidate for the NADPH regeneration further kinetic assays will be applied, also further generations can be constructed using the active mutant as a template to improve the efficiency of the $NADP^+$-dependent FDH for industrial usage.

Candida methylica FORMAT DEHİDROGENAZ'IN NAD^+ / $NADP^+$ SPESİFİTESİ İÇİN KRİTİK BÖLGELERİN BÖLGE SATURASYON MUTAGENEZİ

ÖZET

Metilotrofik organizmalarda metanol metabolizmasının son adımını oluşturan NAD^+-bağımlı format dehidrogenaz enzimi (EC 1.2.1.2, FDH), format iyonunun karbondiokside dönüşümünü katalizlerken, NAD^+ molekülünün NADH'e indirgenmesini sağlamaktadır. Doğada bulunan FDH'lerin çoğunluğunun NAD^+ koenzimine yüksek ilgisinin bulunması, FDH'in NADH rejenerasyonunda yaygın olarak kullanılmasını sağlarken, NADPH rejenerasyonunda kullanımını kısıtlamaktadır. Bunun yanında, FDH'in $NADP^+$ koenzimi ile NADPH rejenerasyonunda kullanımı, düşük maliyeti, termodinamik kararlılığı ve son ürün olan karbondioksidin ortamdan kolay uzaklaştırılması gibi avantajlar sağlamaktadır. Bu amaçla, FDH'in sınırlı koenzim spesifitesinin aşılması için farklı protein mühendisliği çalışmaları gerçekleştirilmiştir. Gerçekleştirilen rasyonel dizayn uygulamaları sonucunda geliştirilen mutantlar tamamen $NADP^+$ spesifitesi gösterememişlerdir.

Bu projede yönlendirilmiş evrim tekniği uygulamalarından biri olan bölge saturasyon mutagenez tekniği kullanılarak *Candida methylica* FDH'in $NADP^+$ spesifisitesinin attırılması amaçlanmıştır. Öncelikle, koenzim bağlama bölgesinde, koenzim spesifitesinden sorumlu amino asitler, *Pseodomonas. sp.*101 ve *Candida boidinii* FDH kristal yapıları baz alınarak, Insight II (Accelrys) programı ile oluşturulan *cm*FDH homoloji modeli yardımıyla belirlenmiştir. Belirlenen D195, Y196 ve Q197 bölgelerine ait dejenere primerler ile uygulanan bölge saturasyon mutagenez çalışması sonucunda her bölge için mutant kütüphaneleri oluşturulmuştur. İstenen özelliğe sahip mutantlar kolorimetrik tarama metoduyla belirlenmiştir. Taranan yaklaşık 400 koloni arasında $NADP^+$ koenzimine aktivite gösteren 27 aday mutant belirlenmiştir. Bu sonuçlar, belirlenen bölgelerin $NAD(P)^+$ spesifitesinde önemli bölgeler olduğunu göstermiştir, ayrıca $NADP^+$-spesifik *cm*FDH geliştirilmesinde umut verici olmuştur. Çalışmanın ileriki aşamalarında, en aktif ve etkili mutantın belirlenmesi için, detaylı kinetik çalışmalr uygulanacaktı. Ayrıca endüstriyel kullanımda daha etkili $NADP^+$-spesifik FDH oluşturma amacıyla, aktif mutant kalıp kullanılarak farklı jenerasyonlar geliştirilebilir.

1. INTRODUCTION

By the increasing energy consumption and because of global warming green processes using biocatalyst become an alternative way over traditional chemical catalysts in medicine, chemical industry, food processing and in agriculture, since enzymes have many advantages such as; biodegradability, high selectivity like stereoselectivity, regioselectivity, and chemoselectivity and they produce enantiomerically pure products. Recently, several industrial enzymes have been used for the synthesis of nearly 500 commercial products in textile, pulp and paper and detergent industries (65%), in food processing (25%), and in animal feed supplements (10%) [16, 19].

Industrial processes often require extreme conditions like high temperature, pressure and pH which require a large amount of energy to achieve. However, enzymes which are naturally occurring biocatalysts have some limitations for industrial usage such as low stability and activity at such extreme conditions. In addition, most of the enzymes have limited substrate and coenzyme specificity and low kcat. Several approaches such as nanotechnology, metabolic engineering, cellular membrane engineering, and protein engineering are applied to overcome the limitations and to increase the applications of the biocatalysts [9, 14, 19].

Protein engineering approaches; rational design, directed evolution and finally combination methods have been developed to overcome these problems and to optimize enzymes for specific industrial applications.

Rational protein-design is the earliest approach which needs structural information and function relationship of the enzyme. In order to improve a specific property of the enzyme, in rational design it is important to identify the specific residues which are responsible for the desired property and this is provided by crystal structure data or data obtained from homology modeling. However, the complexity of the structure/function relationship in enzymes has become a factor limiting the application of rational design. Unlike rational design, directed evolution, which is the most widely used approach, does not require a detailed information of enzyme

structure and structure/function relation. Simply, directed evolution relies on the Darwinian principles of mutation and selection, which consists of the low frequency introduction of randomly distributed mutations by random mutagenesis or gene recombination, followed by selection of the mutant proteins with the desired properties. In directed evolution, it is important to develop a high-throughput screening methodology for the selection of mutants with desired properties, and thus this necessity limits the application of the approach. Recently, combination of the previous approaches has been used more efficiently in modifying industrial biocatalysts. Combination of random methods by using rational modification elements could successfully by-pass certain limitations of both directed evolution and rational design. Site-saturation mutagenesis (a 'semi-rational' approach, which involves all 20 amino asid possibilities randomly at specific, predetermined residues) which takes advantages of both rational design and directed evolution such as applying random mutagenesis at targeted sites in order to increase the probability of beneficial properties [9, 16, 27].

Several synthetically useful enzymes, especially oxidoreductases, are underutilized because of the requirement of expensive reduced nicotinamides (NADH or NADPH) as cofactors. Therefore, regeneration of reduced nicotinamide coenzymes (NAD(P)H) become very crucial for the synthesis of chiral compounds in chemical industry. Because of the highly costs of these coenzymes various methods such as chemical, electrochemical, photochemical and enzymatic methods have been developed for the regeneration of reduced (NAD(P)H) coenzymes. Currently, NAD(P)H regeneration is widely based on enzymatic methods especially dehydrogenase systems. Formate, glucose and phosphite dehydrogenase systems have been applied for the coenzyme regeneration, of them formate dehydrogenase (FDH) is the best and most widely used system for enzymatic NAD(P)H regeneration, because of the advantages of FDH. NAD^+-dependent formate dehydrogenase (EC 1.2.1.2, FDH) plays an important role in the final step of methanol oxidation pathway in methylotrophs, which catalyzes the oxidation of formate anion into carbon dioxide, coupled with reduction of NAD^+ to NADH. Because of the reaction's simplicity, availability and low cost, a favorable thermodynamic equilibrium and the inertness of CO_2, FDH becomes a potential NAD(P)H regeneration system. FDH from *Candida boidinii* is the only enzymatic

NAD(P)H regeneration system which is used industrially (tert-L-leucine production with FDH for NADH regeneration in Degussa Company). However, naturally occurring FDH's are highly specific to formate and NAD^+, so FDHs are not efficient enough for NADPH regeneration. In order to overcome limited coenzyme specifity of FDH, many attempts using a rational design approach have been made recently. However, all these mutants bind $NADP^+$ very weakly and still show activity with NAD^+ [17, 37, 38].

1.1 Protein Engineering

For thousands of years in food production via fermentation, microorganisms and their enzymes have been used naturally. Recently, biocatalysts (whole cells and enzymes) have become a great potential in various applications such as medical therapies, chemical, food or textile industries. However, the use of naturally occurring biocatalysts is hindered by the low stability or activity and lack of substrate or coenzyme specificity in extreme industrial conditions. To overcome the limitation of the natural biotocatalysts, protein engineering tools have been used to enhance the performance of the enzymes under non-natural industrial environments. Two distinct strategies of protein engineering, namely; rational design and directed evolution are currently available to improve stability or activity, limit substrate specificity and alter coenzyme specificity of the enzymes [27].

In general, rational design is based on the structural information and the function/structure relation of the enzyme and changes at the predetermined residues are carried out by site-directed mutagenesis. In contrast, directed evolution does not require structural information of the enzymes. The approach is based on natural evolution and selection. Obtaining the desired property of the enzyme is provided through a high-troughput screening after random mutagenesis and/or gene recombination of the enzyme. Creation of mutant libraries by various directed evolution tools and developing a selection or screening methodology are very critical points of the directed evolution approach. Both of the strategies have a great potential in enzyme optimization for industrial applications and have distinct advantages, however, they also have limitations. Although, rational design enhances our basic knowledge about enzyme structural and catalytic mechanisms, the complexity of the enzyme structure/function relationship limits the application. On

the other hand, applications of directed evolution are hindered by creating inactive enzymes via random mutagenesis or gene recombination and developing an inefficient screening method. In addition, combination strategies which utilize both rational design and directed evolution could successfully by-pass the limitations of these strategies and enhance the properties of an enzyme [6, 27, 40].

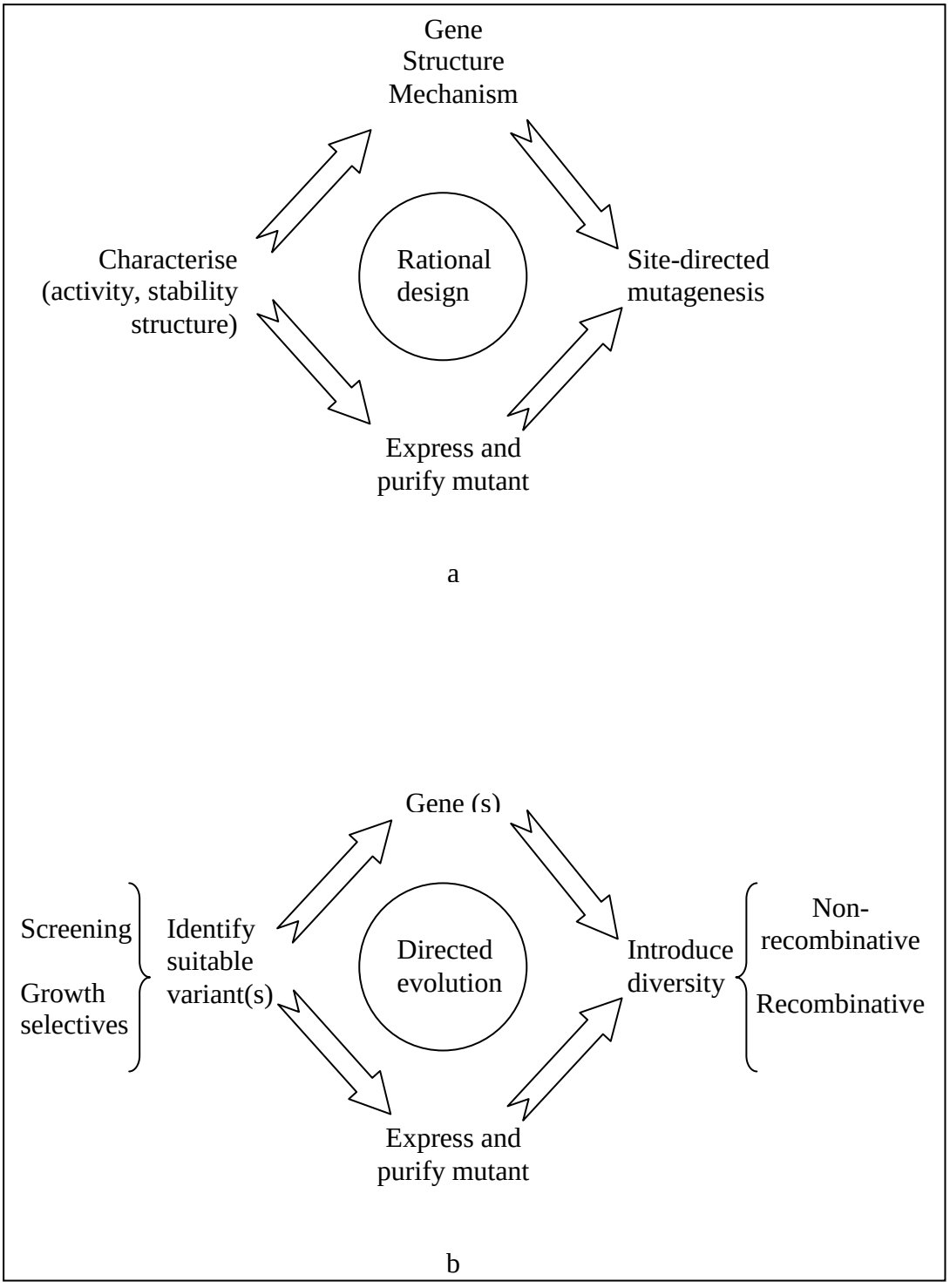

Figure 1.1: General sheme of the protein engineering strategies. (a) Rational Design, (b) Directed Evolution [40].

1.1.1 Rational design

1.1.1.1 Site-directed mutagenesis

Rational design via site-directed mutagenesis is the earliest and most widely used protein engineering approach, which is based on great knowledge of structure/function relationship of the enzyme.

In order to design the new properties or even new proteins, site-directed mutagenesis strategy aims to understand protein structure and function through the 3D protein structure or the homology modeling data. Several successes have been reported in the past, mostly about enzyme specificity and activity by site-directed mutagenesis via active site substitution [5,40]. Another site-directed approach, which is based on structural homology, has been used successfully to alter enzyme mechanism or substrate/cofactor specificity. For instance, introducing Ser-His-Asp into peptidyl-prolyl isomerase resulted in an efficient proline-specific endopeptidase, [29] and superoxide dismutase, which is already one of the fastest known enzymes in nature [13], become faster and also coenzyme specificities of both isocitrate and isopropylmalate dehydrogenases altered successfully by rational design via site-directed mutagenesis [7,8]. Although, site-directed mutagenesis displays great efficiency for the determination of the enzymes' kinetic and functional properties, confirmation of the mutagenesis and purification of the mutant enzymes are required and for each round of the mutagenesis the approach might be expensive and impractical [6]. On the other hand, because of poorly understanding of the relationship between enzyme structure and function, site-directed mutagenesis does not always give the desired outcome [40].

By the increasing knowledge of 3D protein structure data and the development of new protein modeling tools, rational design will become a powerful complementary approach to directed evolution, so semi-rational approaches become more efficient [6,27].

1.1.2 Directed evolution

Directed molecular evolution or *in vitro* evolution, is a general term used to describe the general strategy of mimicking natural evolution in the laboratory and consists of

various techniques for generation of protein mutants (variants) and selection of desirable functions [27, 43, 44]. Since it was first described in the 1970s, directed evolution has come out as a powerful technology and found a wide range of applications in industry, academia, and medicine. In the early studies of "directed evolution", *in vitro* evolution of nucleic acids were carried out [24]. However, several decades later directed evolution concept was applied for *in vitro* engineering of proteins at the molecular level. More recently, directed evolution strategies have been applied for more complex subjects like metabolic pathways, viruses, or bacterial genomes. The procedure begins by determining a target biomolecule, metabolic pathway, or organism, and a desired phenotypic goal. A diversity of mutants is created through the methods that mimic the strategies of traditional evolution, such as random mutagenesis and/or gene recombination. A high-throughput screening or selection method is used to identify the individuals with the desired properties among the library; if necessary, selected mutants can be used as parents in the second round of the cycle. Because, the process is repeated until the phenotypic goal is achieved. Unlike natural evolution, with which the diversity can be achieved after thousands or even millions of years, using directed evolution a meaningful diversity can be created and selected in a much shorter time, like in several weeks.

However, for successful directed evolution there are some requirements described as follows:

i. The desired function must be physically possible.

ii. The function must also be biologically or evolutionarily feasible. In practice, this means that there exists a mutational pathway to get from here to there through ever-improving variants.

iii. It must be possible to create libraries of mutants complex enough to contain rare beneficial mutations.

iv. A rapid screen or selection that reflects the desired function is required[4]

A range of strategies for the introduction of diversity into the starting gene(s) are available, and these can be broadly divided into two classes; non-recombinative design and recombinative design, and can range from creating libraries with as few as 200 variants to many tens of thousands of variants. Non-recombinative method

consist of *in vitro* random mutagenesis which have been developed to generate substitutions, deletions, and insertions. Recombinative method generally consists of homologous or non-homologous gene recombination, which refers to the exchange of blocks of genetic material among two or more DNA strands [27, 40, 43].

1.1.2.1 Non- recombinative design

Non-recombinative design consists of several random mutagenesis strategies, which are relatively simple and popular methods for generating molecular diversity via point mutations, and they have been developed to generate substitutions, deletions, and insertions.

Early methods involved creation point mutations in a target gene through damaging the DNA strand, for example by treatment with chemical mutagens or by ultraviolet irradiation. These methods are inefficient, because they can cause substantial cell damage if performed *in vivo*. Random base pair substitutions can also be generated by error-prone PCR which takes advantages of the fallibility of DNA polymerase, is one of the simplest and most popular directed evolution tools. Also, it is possible to control the rate of mutagenesis in this method.

Error-prone PCR relies on the misincorporation of nucleotides by DNA polymerase to generate point mutations in a gene sequence. In order to improve the efficiency, manganese ion can be added and also nucleotide analogs or "mutagenic polymerases" can be used in the PCR reaction. Because of the relative simplicity and versatility of error-prone PCR, it has become the most widely used mutagenesis strategy, but it has some limitations. The method is limited in its ability to create diversity at the protein level, since DNA polymerases used in PCR reactions have mutational biases that limit diversity. *Taq* polymerase and Mutazyme (Stratagene, La Jolla, California) will preferentially induce mutations at AT base pairs over GC base pairs, this tend to preserve the characteristics of the original residue. Also the mutation rate is low, only 1-3 mutations per 1000 base pairs in general. This may result in low mutation frequencies, limited diversity, and low product yield.

Similarly, mutator strains of *E. coli* exploit defective DNA repair machinery and also create random point mutations. Random insertion and deletion (RID) can also be used for the diversity generation and to modify proteins [27, 40, 41, 43].

The limitations of error-prone PCR mutagenesis may be overcome by saturation mutagenesis or sequence saturation mutagenesis (SeSaM) methods. Saturation mutagenesis involves the substitution of all possible amino acids randomly at a predetermined residue or continuous series of residues in the protein of interest. SeSaM, a recently described method, which is able to randomize a DNA sequence at every nucleotide position through use of a universal base [41, 44]. Various types of random mutagenesis methods are compared in Table 1.1.

Table1.1: Comparison of non-recombinative methods [27]

Method	Advantages	Disadvantages
Chemical Mutagenesis	Simplicity	Accumulates deleterious mutations Low mutation level Low efficiency Limited amino acid substitutions Cannot control mutation rate
Mutator strains	Simplicity	Low mutation level Accumulates deleterious mutations Progeny must be transferred to DNA repair-competent strain for screening Limited amino acid substitutions Cannot control mutation rate
Error-prone PCR	Simplicity	Accumulates deleterious mutations Limited amino acid substitutions Polymerase bias
Saturation Mutagenesis	Simplicity Mutate specific site(s) in a gene Access all 20 amino acid	Limited diversity generation Gene sequence required
Sequence saturation Mutagenesis (SeSAM)	Overcomes polymerase bias Target a specific nucleotide in a sequence	Small fragments not mutagenized Four PCR reaction needed to remove bias Limited amino acid substitutions
Random insertion/ Diletion (RID)	Flexible Insert or remove an aminoacid randomly Acess all 20 amino acid	Point mutations may ocur Time-consuming and technically challenging

1.1.2.2 Recombinative design

Recombinative design based on homologous or non homologous gene recombination refers to the exchange of blocks of genetic material among two or more DNA strands, and is often considered the "sexual" component of evolution. Gene recombination, which can be divided mainly into two groups refer to; homologous recombination, where recombination occurs between two genes with high sequence identity, and non-homologous recombination, where recombination occurs between two DNA sequences with little or no sequence identity.

The first *in vitro* homologous recombination method, DNA shuffling, was introduced by Stemmer in 1994 [36]. Since then, various other recombination methods have been developed.

Homologous recombination strategy mimics the "sexual" recombination of genetic material that rearranges maternal and paternal chromosomes in germ cell DNA. Such recombination increases the genetic variation among a population, through this idea, diversity generation is targeted by gene recombination methods.

Despite the advantages of non-recombinant methods for variant library production, the most significant changes in enzyme function have been created using recombinative methods. DNA shuffling is still the most popular method of recombining DNA. Shortly, DNA shuffling method involves the digestion of the source DNA using DNase I into random fragments, followed by reassembly of those fragments into a full-length gene by a primerless PCR, and final standard PCR reaction for the amplification of the small amount of full length gene in the presence of flanking primers. The fragmentation and reassembly processes cause introducing point mutations and these mutations add to the diversity of the mutant library [27,40].

Several years after the introduction of DNA shuffling, the method was adapted to the recombination of a family of related genes from various species. This method, called family shuffling, based on DNA shuffling method which is applied to a group of naturally occurring homologous genes [10]. To date, various homologous recombination methods have been developed, as well as non-homologous recombination strategies have been improved based on DNA shuffling strategy.

These various homologous and non-homologous recombination methods are compared in Table 1.2. & Table 1.3.

Table1.2: Comparison of homologous recombination methods [27]

Method	Advantages	Disadvantages
DNA shuffling	Robust, flexible Back-crossing to parent removes non-essential mutations	Biased to crossovers in high homology regions Low crossover rate High percentage of parent
Family shuffling	Exploit natural diversity Accelerates functional enzyme improvement	Biased to crossovers in high homology regions Need high sequencehomology in the gene family High percentage of parent
Family shuffling using restriction endonucleases	Lower representation of parent in a library	Point mutations Low crossover rate
DOGS	Reduced parent genes in shuffled library Lower homology required Can bias representation of parent in library	Point mutations Frameshifts may occur Relatively low crossover rate
RACHITT	No parent genes in a shuffled library Higher rate of recombination Recombine genes of low sequence homology	Complex Requires synthesis and fragmentation of single-stranded complement DNA
RPR	Compatible with ssDNA DNaseI-independent Removes sequence bias Independent of template length Less parent DNA needed	Need gene sequence Biased point mutations also occur
StEP	Simplicity	Need high homology Low crossover rate Need tight control of PCR
Synthetic shuffling	Greater flexibility Increased diversity	Chemical synthesis of many degenerate oligonucleotides
Genome shuffling	Improve complex, poorly understood phenotypes Adapt to multiple phenotypic goals	Possibility of novel antibiotic resistance or pathogenicity Genome flexibility restricted by metabolic network rigidity

Table 1.3: Comparison of non-homologous recombination methods [27]

Method	Advantages	Disadvantages
Exon shuffling	Preserves exon function	Requires known intron-exon organization of target gene Limited diversity
ITCHY	Eliminate recombination bias Structural knowledge not needed	Limited totwo parents Significant fraction of progeny out-of-frame Complex, labor-intensive
THIO-ITCHY	Same advantages as ITCHY Combines recombination and random mutagenesis Simplified ITCHY method	Same disadvantages as ITCHY Incorporated dNTP analogs may complicate further experimentation
SCRATCHY	Eliminate recombination bias Structural knowledge not needed	Limited to two parents Significant fraction of progeny out-of-frame Complex, labor-intensive
DHR	High recombination rate Eliminate recombination bias	Synthesize numerous complementary oligonucleotides Gene seuence needed
RM-PCR	Unbiased incorporation of variable size DNA fragments	Frame shifts may ocur Mutants may be longer or shorter than expected
SHIPREC	Crossovers occur at structurally related sites	Limited to two parents Single crossover per gene
SISDC	Recombines fragments without bias Ligates fragments in a desired order	Gene sequence needed Must engineer endonuclease sites into parent genes Must synthesize numerous oligonucleotide primers
YLBS	Recombines variable size DNA fragments Shuffles large fragments such as exons or domains	Non-stoichiometric incorporation of DNA fragments Frame shifts may occur Low product recovery

1.1.3 Semi-rational design

Practical experience shows that directed evolution and rational design can produce remarkable changes in improving biocatalysts. However, these methods have their limits as discussed above. In many cases, a combination of rational design to create the needed structure or function and its improvement by random techniques is a better approach [21].

Recently, combination of random methods of directed evolution with elements of rational enzyme strategies can successfully by-pass certain limitations of both directed evolution and rational design (Table 1.4). In semi-rational approaches, pre-determined specific residues through the basic structural or functional knowledge are randomized using directed evolution tools, especially saturation mutagenesis, create 'smarter' libraries that gave positive results [9].

For example, to improve the synthetic capacity of γ-humulene synthase to produce different sesquiterpenes saturation mutagenesis with systematic recombination approach inside or near the active site have been applied succefully [42].

Table 1.4: Comparison of the strategies of the protein engineering.

	RATIONAL DESIGN	DIRECTED EVOLUTION	SEMI-RATIONAL DESIGN
Knowledge of structure	Required	Not-required	Required
Knowledge of mechanism	Required	Not-required	Required
Screening and selection method	Not-required	Required	Required
Sensitive enzyme assay	Required	Not-required	Not-required

1.1.3.1 SSM: Site saturation mutagenesis

Site saturation mutagenesis (SSM) technology is a unique method for rapid laboratory evolution of proteins which enables to create a library of mutants

containing all possible amino acid (20 naturally occurring amino acids) changes at one or more pre-determined target positions in a gene sequence.

The approach is applied at the genetic level through the use of standard DNA amplification by degenerate primer sets, containing either 32 or 64 codon variants for each amino acid residue and finally a library is constructed that contains all amino acid possibilities at the determined sites.

In combination with high-throughput screening methods, the researchers have successfully used saturation mutagenesis to improve enzymatic properties such as thermostability [23,25].

Randomizing (mutated randomly) active-site residues by using site-saturation mutagenesis achieved high β- galactosidase activity (180-fold increase) in comparison to DNA shuffling (10-fold increase) approach and gave greater substrate specificity [26].

For developing nitrilase as a process-scale enantioselective biocatalyst is one of the applications of this method. Mutagenesis and screening resulted in a nitrilase variant with high enantiomeric excess (*ee*) at high substrate concentrations [12].

To improve catalytic properties of P2Ox from *Trametes multicolor* and its substrate specificity, the semi-rational approach was selected for the engineering. Screening after the saturation mutagenesis of the active site residues showed catalytic constants increased by up to 5.7-fold for both the sugar substrates (d-glucose and d-galactose) [35].

Also, combination of site-saturation and random mutagenesis approaches have improved the properties of the target biocatalysts. Several studies involving combination of error-prone PCR or DNA shuffling with site- saturation mutagenesis have created succesfull variants [9]. In the absence of structural/functional information, these randomizing strategies can enable primary information for the targeted strategies; after several rounds of random mutagenesis or DNA shuffling site-saturation mutagenesis can easily be applied based on the knowledge. Therefore, the structure-based site-saturation mutagenesis becomes a power by enhancing our capacity for rational design, while using advantages of random mutagenesis. And also, this approach becomes a powerfull tool for the protein engineering that take

advantages of both strategies and can by-pass the limitations of both strategies [9, 27, 40].

1.2 Selection and Screening

In protein engineering, developing an effective protocol to screen the target enzyme libraries with the desired properties is more critical than creation the diversity. Generally, directed evolution tools and semi-rational approaches need well designed high throughput screening methodology or direct selection methods.
Screens for enzyme activity usually operate via detection of optical adsorption or fluorescence. Some of the screening methods can apply directly on agar plates, where changes of color are observed by direct visualization, and this screening process can be relatively simple.

However, most screens involve transfer of the colonies into multi-well microplates which contain culture medium, growing the cells until stationary phase, and inducing protein expression. Cell lysis step changes according to the enzyme being intracellular or extracellular. If the enzyme is intracellular, extra lysis step has to be carried out. Finally, enzymatic activity is assayed with microplate reader. Through this informations, it is important to develop a sensitive, effective, and easy screening or selection method suitable for the target enzyme activity.

Screening may be achieved by genetic selection, where the transformed organism can only survive if the desired activity is present. The system is very efficient when it is set up for searching for a protein function which provides a growth advantage to the microorganism.

However, majority of applications of directed evolution are usually impossible using selection methods. Despite genetic selections have been used to great effect, many researchers concentrate on developing screening protocols, When selection is not possible, every member present in the library must be physically separated and individually assayed for function. This can be done either in agar plates or in microtiter plates, using sensitive detection methods for catalysis adapted to high-throughput.

Fluorogenic or chromogenic assays are widely used approaches for screening the properties of the target enzyme. These colorimetric or fluorometric methods are

notably the simplest, the most reliable, and the easiest to convert to high-throughput format for screening. Where products of the reaction are not directly colored or fluorescent, indirect assays can be used for screening of the partial activity. Traditional assays for purified enzymes can be scaled down for use in 96-well microtitre plates. Colonies can be grown directly in microtitre plates and lysed if necessary. Often, indirect screen which is a rapid primary screen, can be used to eliminate clones with zero activity, then combined with a more sensitive secondary screen.

Fluorometric and colorimetric screens are the most sensitive and convenient approaches to use; however, they are not available for most enzymes. Also, instrumental approaches are avaliable for the library screening. The generic screening tools include HPLC, gas chromatography (GC), NMR, mass spectrometry (MS), and capillary array electrophoresis (CE) can be given as an example. The development of these assays can be achieved much faster than the chromogenic one, and they can also yield a much greater quantity of useful information; however, these assays are expensive [2, 3, 39, 40].

1.2.1 Colorimetric assay for NAD(P)H generation

Commonly, dehydrogenase activity is measured directly by the production of NAD(P)H at 340nm, but this procedure for screening of large libraries with cell lysates in microplates is not suitable. The high costs of the microplates for screening in the UV-range, the low reproducibility of the measurements, can limit the efficiency of the approach.

Therefore, colorimetric assay for the NAD(P)H generation promises to be generally applicable for measuring the activity of dehydrogenases.

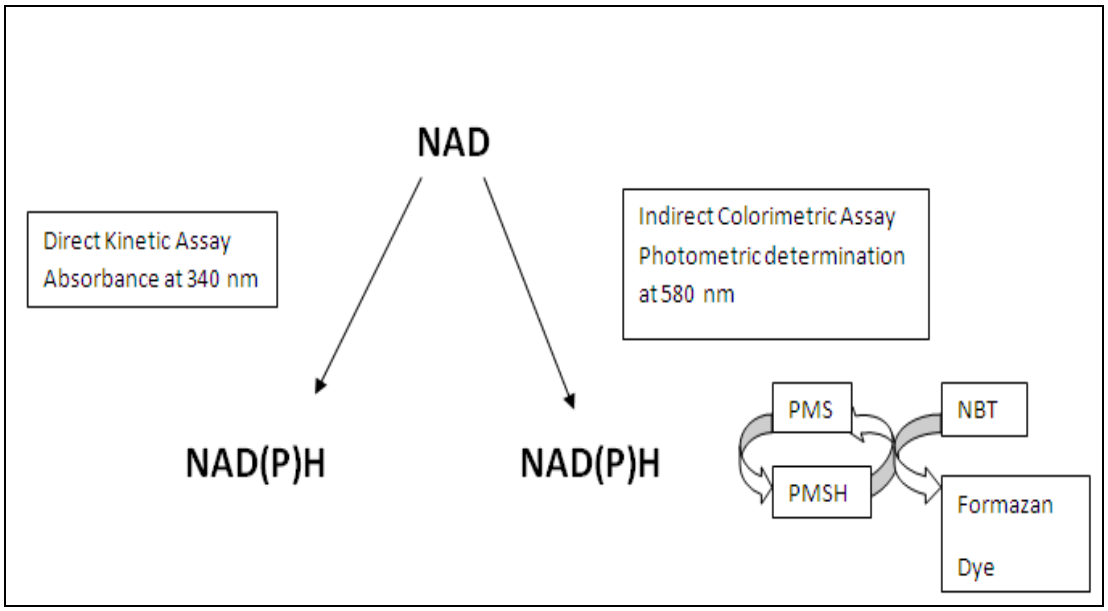

Figure 1.2: Direct and indirect determination of FDH dehydrogenase activity.

Colorimetric solid or liquid phase assays monitor NAD(P)H production indirectly by reduction of tetrazolium, e.g., nitroblue tetrazolium (NBT), salts to formazan dyes.

Colorimetric assay is based on the reduction of nitroblue tetrazolium (NBT) to soluble formazan in the presence of phenazine methosulfate (PMS) which reacts with the NAD(P)H produced by dehydrogenases (Figure 1. 3.)

If there is an active dehydrogenase in the reaction, NAD(P), and PMS, NBT is reduced to blue-purple formazan that, when scanned spectrophotometrically in the range of 400-700 nm, shows maximum absorption at ≈ 560 nm. Finally, the activity can be measured through blue-purple formazan at 580nm [3, 11, 22].

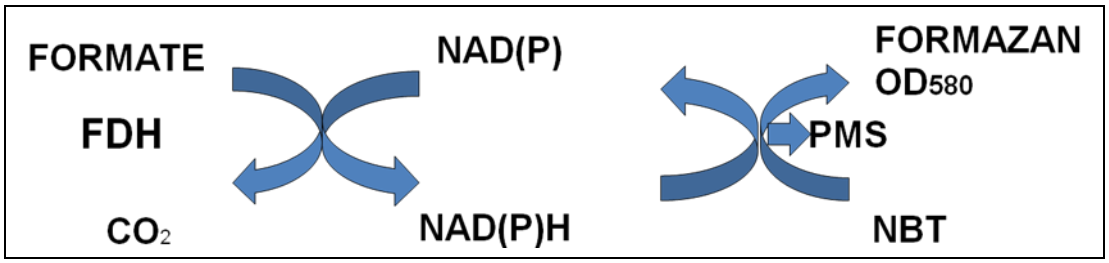

Figure 1.3: General scheme of the colorimetric assay for NAD(P)H generation of formate dehydrogenase.

1.3 NAD^+-Dependent Formate Dehydrogenase

FDHs belong to the superfamily of D-specific 2- hydroxy acid dehydrogenases, which combines several groups of enzymes strongly varied in quaternary structure, presence and type of prosthetic group, and also in substrate specificity. There are three families of FDH. Two of them are complex and use heavy metals such as molybdenium, selenium, iron, etc. The first class of FDH; complex non-NAD^+-dependent FDH has complex subunit structure and requires a wide variety of cofactors and metals. The second class of FDH; complex, soluble NAD^+-dependent FDH uses NAD^+ as a cofactor, but has a complex subunit structure like first class of FDH. And the third one which is the simplest and is called NAD^+-dependent formate dehydrogenase; (EC 1.2.1.2, FDH), only requires NAD^+ as a coenzyme and does not contain any prosthetic groups or metal ions. NAD^+-dependentFDH, plays an important role in the terminal step of the catabolism of C1 compounds in methylotrophs, which catalyzes the oxidation of formate anion into carbon dioxide, coupled with reduction of NAD^+ to NADH (Figure 1.4.)

The enzyme was first discovered in pea seeds more than 60 years ago, and the studies began in 1970s, and mostly founds in yeasts, bacteria, plants, and fungi. FDH, plays an important role in the energy supply of methylotrophic microorganisms and in the stress response of plants. In plants, enzyme localizes in mitochondria and its biosynthesis increases under stressful conditions.

The clear simplicity of the substrate and the reaction makes FDH most suitable model for investigating the general mechanism of catalysis involving hydride ion transfer. In addition, FDH is one of the most promising candidates for the development of so-called coenzyme regeneration systems [1, 28, 31, 37, 38].

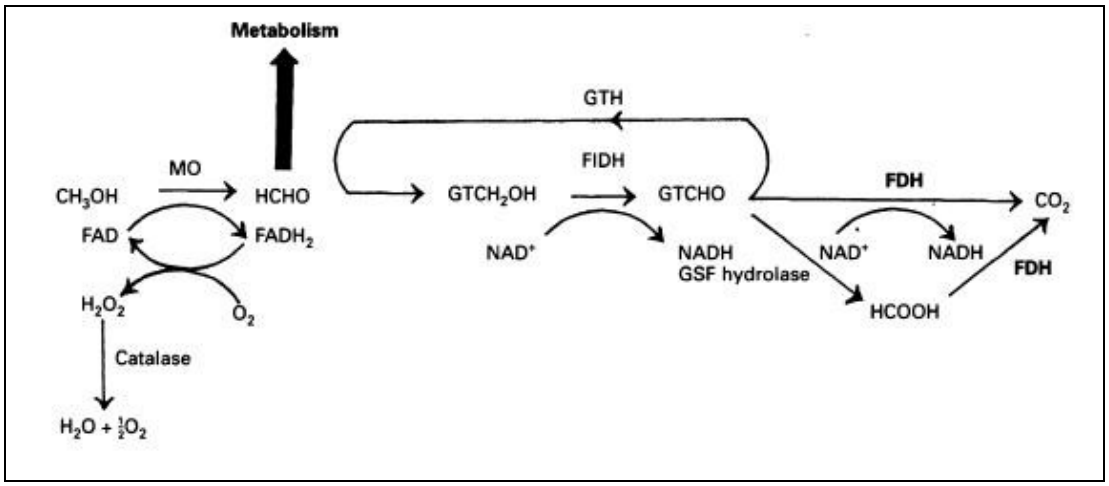

Figure 1.4: General scheme of methanol metabolism in yeasts [28].

1.3.1 Catalytic properties of FDH

Generally, FDHs follow Bi-Bi two-substrate order reaction with NAD^+ as the first substrate. Both substrate and co-enzyme sites are pre-existed, and binding of one of the substrate increases the affinity of the other by 3,5 folds. The catalytic mechanism of this enzyme is included by a direct transfer of hydride ion from the substrate onto the C4-atom of the nicotinamide moiety of NAD^+, which are present in reactions catalyzed by other related dehydrogenases. Majority of FDHs display Michaelis-type kinetics and independent functioning of the active centres. Recent studies of FDH from various organisms show similar kinetic properties, all the enzymes have similar Km values for formate (3-10 mM) and NAD^+ (35-90 μM), and the recombinant FDHs are slightly different from the native enzymes. The majority of NAD^+-dependent FDHs are highly specific to NAD^+ and do not utilize $NADP^+$ as a coenzyme, with only one exception of *pseudomonas* FDH displays dual coenzyme specificity and shows highly activity againts $NADP^+$. FDHs can catalyse pH values between 6.0 and 9.0, and they can keep half of their activity between 50-60 °C with a few exceptions. Though they can work in a fairly wide range of pH, FDH is vulnerable to inactivation at higher temperatures [28, 37].

1.3.2 Structural properties of FDH

NAD^+-dependent FDHs do not contain any prosthetic groups or metal ions. The molecular masses of the enzymes from eukaryotic organisms and some methylotrophic bacteria range from 70 to 100 kDa [28]. NAD-dependent FDHs generally form homodimers and are composed of two chemically identical subunits,

and are highly specific to both formate and NAD^+. Subunits consist of two domains, ‘‘NAD binding domain’’ and ‘‘catalytic domain’’ [31].

All FDHs can be divided into two groups; FDHs from bacteria and plants is the first group, and FDHs from yeasts and fungi is the other group. FDH is a highly conservative enzyme. The homology between enzymes of the same group is nearly 80-85% and between two enzymes from the different groups is 50-55% and more. The comparison of recently known (completely and partially) FDH sequences from various sources showed that nearly 20% of all residues (71 residues) are conserved. In addition, the catalytic amino acids, as well as the amino acids that contribute to the structural stability, are almost totally conserved (sequence homology of approximately 95%) [37].

Residues taking a role in the active site and co-enzyme binding sites are strictly conserved.

Catalytically important amino acid residues of *Pseudomonas* sp. 101 FDH (*ps*FDH) are Pro97, Phe98, Ile122, Asn146, (Ala/Gly)198, Gly200, Gly203, Arg284, Gln313, and His332, in *Candida methylica* FDH(*cm*FDH) Pro77, Phe78, Ile102, Asn118, Gly171, Gly173, Gly176, Arg267, Gln278 and His310 are the residues constructing the active site.

The NAD binding domain (residues N119 to S313) (the residues belongs to *Candida boidinii* FDH (*cb*FDH)) shows a Rossmann fold structure commonly found in members of the dehydrogenase family. The catalytic domain is formed by the remaining residues and has a flavodoxin-like topology. The two domains are linked by two long helices, H6 and H15 [31].

Conservative ‘fingerprint’ sequence GXGXXGX17-18D(E) in the NAD^+ binding domain is specific for the FDH structure. Negatively charged aspartic acid (D195 in *cb*FDH, D221 in *ps*FDH) or glutamic acid at the conserved ‘fingerprint’ sequence is critical for NAD^+-dependent dehydrogenases, and plays an important role in providing the specificity to NAD^+ versus $NADP^+$. Both *cb*FDH and *ps*FDH structures have similar dimer organization. The first glycine in fingerprint sequence is replaced to alanine in *ps*FDH. Also, there is a minor influence between *ps*FDH and *cb*FDH is the additional residue (K189) found in *cb*FDH which is between glycine triplet and the catalytic aspartate. Most of the residues involved in hydrogen bonds to

the cofactor are conserved and occur in similar conformation in *cb*FDH and apo-*ps*FDH. D282 and S313 make contacts to the nicotinamide ring, and R174 binds the phosphate linker in NAD^+, and the adenine ring is likely to interact with H232 and Y196 (the residues are belonging to *cb*FDH). Each of the glycine residues in this fingerprint has a specific role, the first glycine is critical for the tightness of the turn, the second prevents steric hindrance of the dinucleotide and amino acid side chain of the protein backbone at this position, and the third is essential for proper interactions between the strand and helix. The third glycine allows the tight packing of helix H9 onto strand S6, which corresponds to the first strand and its following helix in the cofactor-binding domain, which is in agreement with the general features of the three-glycine pattern of dehydrogenases. (D195 and Q197 in *cb*FDH interact with a phosphate group attached to the O29 of the NAD-ribose, additionally, Y194 and Y196 form a hydrophobic cluster, which could stabilize the adenine ring in a different position and environment [31]. Several studies for the changing coenzyme specifities by using site-directed mutagenesis of FDH from *cm*FDH, *ps*FDH, and *Saccharomyces cerevisiae* FDH (*sc*FDH) have shown the importance of the residues in the NAD^+ specificity [18, 32, 33].

1.3.3 Practical applications of FDH

1.3.3.1 NAD(P)H regeneration

Optically active compounds are very crucial for the chemical and pharmaceutical industry. However, according to prescriptions of the Food and Drug Administration, the optical purity of all chiral compounds used as drugs has to be more than 99%. After these prescriptions, the enzyme applications in pharmaceutical industry are sharply increased. Oxidoreductases, especially, dehydrogenases can be used to produce optically active compounds from nonchiral ones, because dehydrogenases are extremely stereospecific in the transfer of hydride ion between the substrate and coenzyme. Therefore, these enzymes are promising for the production of optically active compounds with very high optical purity (99.9-99.99%). However, dehydrogenases are under utilized because of the requirement of expensive coenzymes NADH and especially of NADPH (>US$ 12,000/kg). Therefore, regeneration of reduced nicotinamide coenzymes (NAD(P)H) become very crucial for the synthesis of chiral compounds in chemical industry. Because of the high costs

of these coenzymes, various methods such as chemical, electrochemical, photochemical and enzymatic methods have been developed for regenerating reduced (NAD(P)H) coenzymes. Currently, NAD(P)H regeneration is widely based on enzymatic methods especially dehydrogenase systems. Formate, glucose and phosphite dehydrogenase systems have been applied for the coenzyme regeneration, of them formate dehydrogenase is the best and the most widely used system for enzymatic NAD(P)H regeneration because of the advantages of FDH [38].

In NAD(P)H regeneration FDH provides all the criteria in the reaction;

i. The reaction by FDH obtain 99–100% yield of the final product, because of the irreversibility of the raction, which provides thermodynamic pressure to shift equilibrium of the main reaction.

ii. FDH exhibits a wide pH-optimum of catalytic activity (6.0– 9.0), thus FDH can be used in combination with any dehydrogenase that has optimum activity in this range.

iii. The substrate of the reaction (formate-ion) is relatively cheap, and the reaction product, CO_2, can be easily removed from the reaction mixture and does not interfere with the purification of the final product.

iv. FDHs from several sources are highly stable enzymes and can be used in the system for a long time.

v. FDHs are inexpensive and available, also methylotrophic bacteria or yeast can provide a high scale enzyme production with a comparatively low production cost.

Therefore, FDH from *Candida boidinii* is the only enzymatic NAD(P)H regeneration system which is used industrially (tert-L-leucine production with FDH for NADH regeneration in Degussa Company).

The reaction catalyzed by FDH is suitable for the system of NADH regeneration. However, all naturally occurring FDHs are highly specific to NAD^+. On the other hand, NADPH regeneration is also critical for the chemical industry. Therefore, protein engineering of FDH with the aim of altering coenzyme specificity ($NADP^+$ versus NAD^+) is become crucial [1, 17, 31, 33, 37, 38]. Recently, several approaches using rational design to alter the specificity of FDH from *Pseudomonas sp*. 101, *Saccharomyces cerevisae* and *Candida methylica* towards $NADP^+$ have yielded

promising results, however FDHs are stil show high NAD^+ activity [18, 32, 33] (Table 1.5).

Table 1.5: Applications for changing coenzyme specifities of FDH's

Aim	Mutation	Result	Ref.
Change coenzyme specificity of *cm*FDH from NAD^+ to $NADP^+$	D195S *cm*FDH	Decrease in coenzyme preference for NAD^+ from 2.5×10^5 to 410	[18]
Change coenzyme specificity of *cb*FDH from NAD^+ to $NADP^+$	D195S D195S/Y196H D195S/Y196H /K356T, *cb*FDH	Activity with NAD^+ and $NADP^+$ 1.5 and 0.083 U/mg, respectively Activity with NAD^+ and $NADP^+$ 1.5 and 0.083 U/mg, respectively Activity with NAD^+ and $NADP^+$ 1.5 and 0.083 U/mg, respectively	[45]
Change coenzyme specificity of *sc*FDH from NAD^+ to $NADP^+$	D196A/Y197R , scFDH	Shift in coenzyme preference for NAD^+ from $> 3 \times 10^9$ to 0.43- 0.67 resulted in $NADP^+$-specific enzyme	[33]
Change coenzyme specificity of *ps*FDH from NAD^+ to $NADP^+$	*ps*FDH T5M9-10	Shift in coenzyme preference for NAD^+ from 2.4×10^3 to 0.29 resulted in $NADP^+$-specific enzyme, K_m^{NADP+} is constant in pH range 6.0-7.0.	[32]

1.4 The Aim of the Research

As explained above, FDH is the best candidate for the NADH regeneration system, because it is stable and it has relatively good activity. FDH from *Candida methylica* (*cm*) was cloned and overproduced at the University of Bristol and purification processes have been improved at Department of Molecular Biology and Genetics of Istanbul Technical University (ITU) giving a much better yield. The one disadvantage of FDH is that it uses only NAD^+ as a coenzyme. It would be also desirable to regenerate NADPH by using $NADP^+$ as a coenzyme. Many attempts

using a rational design approach have been made to change the coenzyme specificity of FDH from NAD^+ to $NADP^+$ but all these mutants bind $NADP^+$ very weakly and still show activity with NAD^+.

Here, we aim to use site saturation mutagenesis which is a technique using a directed evolution approach to the redesign of proteins to improve the K_M of *cm*FDH for $NADP^+$.

2. MATERIALS AND METHODS

2.1 Materials and Laboratory Equipments

Equipments, media, chemicals and enzymes which were used in this study were given in Table 2.1, 2.2, and 2.3.

Table 2.1: Laboratory Equipments

Vortex	Scientific Industries
pH Meter	InoLab
Autoclaves	Tuttnauer 2540ml (Switzerland)
Magnetic Stirrer	Heidolph
Micropipettes	Eppendorf,Gilson
Orbital Shaker Incubator	Biolab-Certomat (Germany)
Microplate Shaker Incubator	Ika
Microfuge	Microfuge 18 Beckman
Centrifuge	Allegra 25R Centrifuge Beckman
UV-Visible Spectrophometer	Shimadzu UV-1601 (Japan)
Microplate Reader	Perkin Elmer
Thermocycler	Biometra
Cycle Sequencer	ABI 3130 Avanti
Electroporator	Eppendorf
Thermomixer	Eppendorf
Deep freezers (-80 °C)	Ultra Low Sanyo
Freezer (-20 °C)	Biomedical Freezer Sanyo
Refrigerator (+4 °C)	Arçelik (Turkey)

Table 2.2: Media

LB (Luria-Bertani) Medium	NaCl	Carlo Erba
	Tryptone	Acumedia
	Yeast Extract	Acumedia
	Agar (for solid media)	Merck
SOC Medium	NaCl	Carlo Erba
	Tryptone	Acumedia
	Yeast Extract	Acumedia
	KCl	Merck
	$MgCl_2$	Merck
	Glucose	Sigma
MM (Magic Media)	*E.coli* Expression Medium	Invitrogen

Table 2.3: Chemicals, enzymes and used kits

Chemicals	Tris-Base	Carlo Erba
	Boric Acid	Merck
	EDTA	BDH Laboratory
	Gelatin	Merck
	NBT	Sigma
	PMS	Sigma
	NADP	Roche
	Formate	Aldrich
	IPTG	Applichem
	Sodium Asetate	Fluka
	Glyserol	Carlo Erba
	Ampicillin	Roche
	BugBuster	Novagen
	Ethanol	Merck
	Agarose	Applichem
	Ethidium Bromide	Fluka
	dNTP	Roche
Chemicals	10X Pfu Buffer	Fermentas
	Tango Buffer	Fermentas
	Marker 3 “ Lamda DNA”	Fermentas

Table 2.3 (continued): Chemicals, enzymes and used kits

Enzymes	Pfu Taq Polymerase	Fermentas
	PstI	Fermentas
	SacI	Fermentas
	DpnI	Roche
Used Kits	High Pure Plasmid Isolation Kit	Roche
	BigDye Terminator V3.1 Cycle Sequencing Kit	Applied Biosystems

2.2 Template DNA

Plasmid DNA included of *Candida methylica* FDH gene, which was previously [46] inserted at the *Sac*I / *Pst*I restriction sites of the pQE-2 vector, and transformed into *Esherichia coli* host cell, was used as a template in this study.

2.3 Methods

2.3.1 Homology modeling

According to the structural data and sequence analysis, the coenzyme-binding domain of FDHs has a classical Rossmann fold. NAD^+-specific oxidoreductases and all FDHs belong to superfamily of D-specific 2-hydroxy acid dehydrogenases involve a conserved `fingerprint' sequence G(A)XGXXG and a conserved aspartic acid residue (Asp221, Asp195 and Asp196 in *ps*FDH, *cm*FDH/*cb*FDH and *sc*FDH respectively). The X-ray data for *ps*FDH and *cb*FDH show that, this conserved Asp residue interacts with the 2'- and 3'-OH groups of adenosine ribose and this residue is a major factor of the specificity for NAD^+[20, 31]. The catalytic Asp is located 18 residues downstream from the Gly residue at the end of the `fingerprint' sequence in yeast FDHs, the bacterial and plant sequences have the conserved Asp as the 17th residue downstream from the end of the `fingerprint'. In addition to the conserved Asp196 residue, tyrosine residue in yeasts (Tyr195 in *cb*FDH and *cm*FDH, Tyr196 in *sc*FDH) is replaced by Arg in bacterial and plant enzymes, and this residue is located at the entrance of the coenzyme binding site and it is the potential residue that prevents $NADP^+$ binding either by unfavourable interactions or by sterically blocking 2'-phosphate group binding. According to these informations, for the determination of the residues that are responsible for the NAD^+ specificity, the

sequences of the *cm*FDH is compared with *ps*FDH and *cb*FDH, and it shows 37% and 97% homology respectively. The comparison of the *cm*FDH and *cb*FDH sequences are shown in Figure 2.1. The conserved glycine triplet is in red color, and the residues responsible for the NAD^+ binding D195, Y196, and Q197 are in green, blue, and yellow color respectively.

```
FDH_Cbo         MKIVLVLYDAGKHAADEEKLYGCTENKLGIANWLKDQGHELITTSDKEGGNSVLDQHIPD 60
FDH_Cme         MKIVLVLYDAGKHAADEEKLYGCTENKLGIANWLKDQGHELITTSDKEGETSELDKHIPD 60
                                                                .. .  .

FDH_Cbo         ADIIITTPFHPAYITKERIDKAKKLKLVVVAGVGSDHIDLDYINQTGKKISVLEVTGSNV 120
FDH_Cme         ADIIITTPFHPAYITKERLDKAKNLKSVVVAGVGSDHIDLDYINQTGKKISVLEVTGSNV 120
                                .       .

FDH_Cbo         VSVAEHVVMTMLVLVRNFVPAHEQIINHDWEVAAIAKDAYDIEGKTIATIGAGRIGYRVL 180
FDH_Cme         VSVAEHVVMTMLVLVRNFVPAHEQIINHDWEVAAIAKDAYDIEGKTIATIGAGRIGYRVL 180
                                                              * *  *

FDH_Cbo         ERLVPFNPKELLYYDYQALPKDAEEKVGARRVENIEELVAQADIVTVNAPLHAGTKGLIN 240
FDH_Cme         ERLLPFNPKELLYYDYQALPKEAEEKVGARRVENIEELVAQADIVTVNAPLHAGTKGLIN 240
                   .          ***

FDH_Cbo         KELLSKFKKGAWLVNTARGAICVAEDVAAALESGQLRGYGGDVWFPQPAPKDHPWRDMRN 300
FDH_Cme         KELLSKFKKGAWLVNTARGAICVAEDVAAALESGQLRGYGGDVWFPQPAPKDHPWRDMRN 300


FDH_Cbo         KYGAGNAMTPHYSGTTLDAQTRYAQGTKNILESFFTGKFDYRPQDIILLNGEYVTKAYGK 360
FDH_Cme         KYGAGNAMTPHYSGTTLDAQTRYAEGTKNILESFFTGKFDYRPQDIILLNGEYVTKAYGK 360
                                        .

FDH_Cbo         HDKK 364
FDH_Cme         HDKK 364
```

Figure 2. 1: Sequence alignments of *cb*FDH and *cm*FDH.

In the coenzyme binding domain, the amino acid residues, which are responsible for the coenzyme specificity was determined by using Insight II (Accelrys) program on a homology model of *cm*FDH based on *ps*FDH (sp.101) and *cb*FDH crystal structure. The determined residues are; D195, Y196, and Q197,which are critical for the NAD^+ binding domain (Figure 2. 2.).

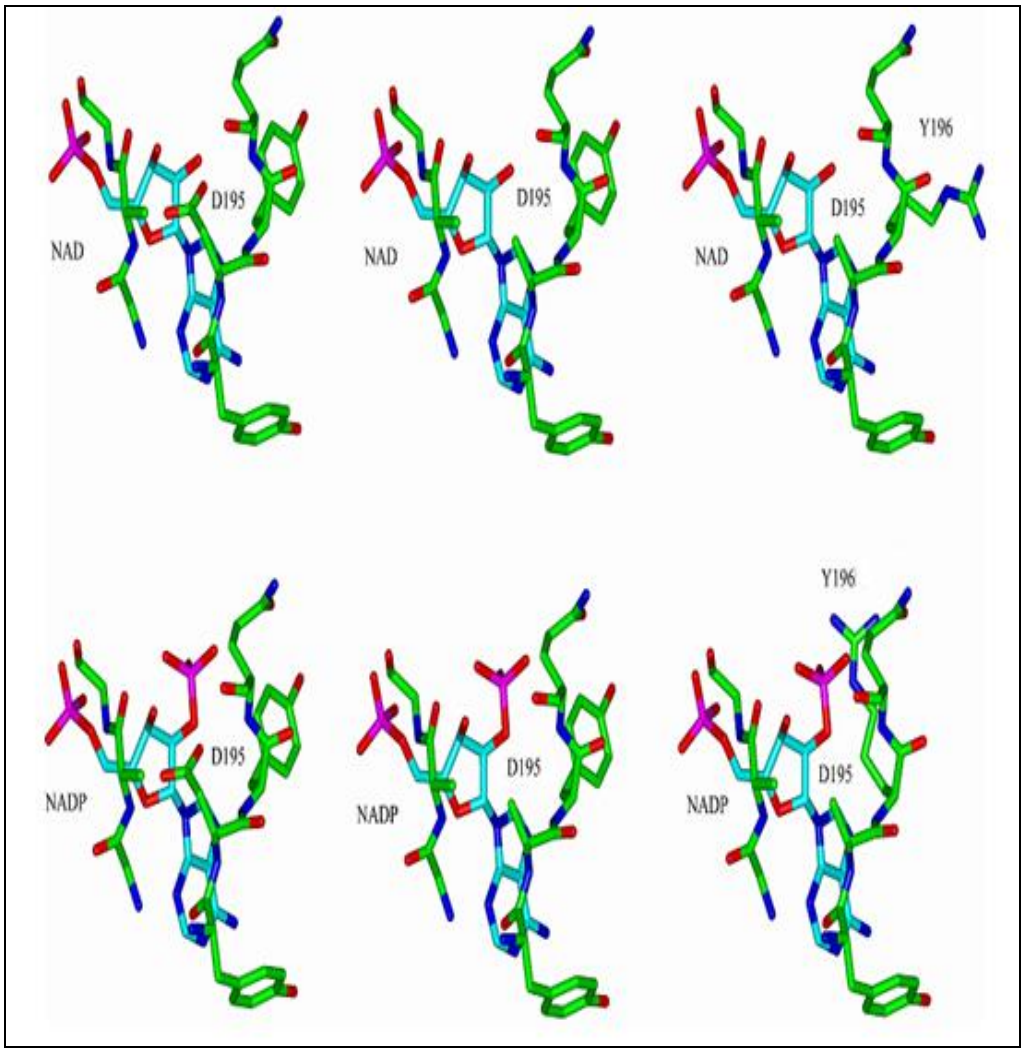

Figure 2.2: Homology model of the determined residues which are responsible for the NAD^+ specificity of *cm*FDH.

2.3.2 Library construction with site-saturation mutagenesis (SSM)

2.3.2.1 Plasmid isolation

For simultaneous purification of plasmid DNA from *E. coli* cultures, which were grown after overnight (≈12-16h) incubation in LB (Luria-Bertani) medium at 37°C, 'High Pure Plasmid Isolation Kit' (Roche) was used. After overnight incubation the cells were harvested, and the supernatant was discarded. Bacterial pellet was resuspended with 250 µl suspension buffer containing RNase, then, 250 µl lysis buffer was added to the homogenate and incubated for 5 min at room temperature (RT). After incubation, lysed solution was treated with 350 µl previously chilled binding buffer, and incubated on ice for 5 min. The mixed solution was then centrifuged at 13,000 x *g* for 10 min at RT. The supernatant was then passed through

a high pure filter tube, allowing selective and efficient binding of plasmid DNA. In this step high pure filter tube was placed on a collection tube and centrifugation was performed at 13,000 x *g* for 1 min. The filter tube was replaced onto the collection tube, after discarding the flow through liquid. The filter tube was then washed with 700 µl wash buffer. After centrifugation for 1 min flow through liquid was discarded, and additional centrifugation for 1 min was applied for the removal of residual wash buffer. Finally, the DNA filter tube was put onto a new, sterile 1,5 ml microcentrifuge tube, and plasmid DNA was eluted with 100 µl elution buffer after centrifugation for 1 min. The eluted plasmid DNA was then stored at -20°C for later analysis.

2.3.2.2 Site-saturation mutagenesis polymerase chain reaction (SSM-PCR)

Primer design

The determined residues (D195, Y196, and Q197) in the NAD-binding domain in *cm*FDH gene were targeted for mutational analysis. For this purpose, four sets of primers were designed for the amplification of the whole plasmid with the degeneracy of the determined sites. Three sets of primers (195F/R, 196F/R, and 197F/R) were designed for single target residue, one set of primer (195 /6 /7 F-R) was designed for triple target residues. For the introduction of the mutagenesis degenerate primers, including all (20) amino acid posibilities were designed (Table 2. 1.). Primer sequences to construct the diversity of the determined sites were designed using the Primer3 software (available online at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi).

Table 2.4: Degenerate oligonucleotide primer sets of the determined residues for the site-saturation mutagenesis PCR

Primers	Sequence
195 F	caaaagaattattatactacNNKtatcaagctttacc
195 R	ggtaaagcttgataMNNgtagtataataattcttttg
196 F	gaattattatactacgatNNKcaagctttacc
196 R	ggtaaagcttgMNNatcgtagtataataattc
197 F	ctacgattatNNKgctttaccaaaagaagc
197 R	gcttcttttggtaaagcMNNataatcgtag
195 /6 /7 F	gaattattatactacNNKNNKNNKgctttaccaaaagaagc
195 /6 /7 R	gcttcttttggtaaagcMNNMNNMNNgtagtataataattc

N= A, T, G, C K= G or T; M= C or A.

SSM-PCR

Site-saturation polymerase chain reaction was used to amplify the whole plasmid DNA for introduction of diversity to predetermined sites of the *cm*FDH by using degenerate primers. For this purpose, the isolated plasmid DNAs of the study were used as templates. The mixture was used for the SSM-PCR. (Table 2. 2.)

Table 2.5: Site-saturation PCR mix

Ingredient	Stock Conct.	Volume	Final Conct.
*Pfu*Taq Buffer	10X	5µl	1X
dNTP mix	10mM	1.5 µl	0.3mM
Forward Primer	10 µM	1 µl	0.2 µM
Reverse Primer	10 µM	1 µl	0.2 µM
Taq Polymerase	5U/ µl	0.5 µl	2.5 U
dH2O	-	40 µl	-
Template DNA	50ng / µl	1 µl	1ng
TOTAL	50 µl		

PCR conditions

Table 2.6: Site-saturation PCR cycle conditions

Cycle Number	Degree	Time	Phase
1	95 °C	2 min	Initial Denaturation
35	95 °C	2 min	Denaturation
	50 °C	2 min	Annealing
	72 °C	8 min	Elongation
1	72 °C	10 min	Final Extension
1	4°C	∞	Final Hold

2.3.2.3 Agarose gel electrophoresis

Due to the size of the PCR products being nearly 6 kb (4800 bp PQE-2 plasmid + 1094 bp *cm*FDH gene), 1% agarose gels were used in this study. For this purpose, agarose gels were composed of 0,40 mg agarose, 2 µl (0,4 µg/ml) of ethidium bromide, and 40 ml of 1X TBE buffer, which is diluted from 10X stock TBE. 5 µl of PCR product was mixed with 1 µl 6X loading dye (Fermentas) and loaded into the wells. Marker 3 ‘‘Lambda DNA / *Eco*RI+ *Hind* III ’’(Fermentas) was used as DNA marker. The gels were run in 1X TBE buffer, at 90V with power supply, for at least 30 minutes. The gels are observed under UV light.

2.3.2.4 *Dpn*I digestion

After amplification of the plasmid DNA, the PCR product was digested with 10 units of *Dpn*I (Roche) for 4 hours at 37°C. *Dpn*I endonuclease recognises and cleaves the methylated adenine sites, and dam^+ strains are suitable for this restriction (Figure 2. 3.). Therefore, restriction of the PCR products provide the elimination of the ancestral dam^+ templates with no mutation. Our templates, were also dam^+, and for further studies with mutated plasmids, the ancestral plasmids with no mutation were eliminated by *Dpn*I digestion.

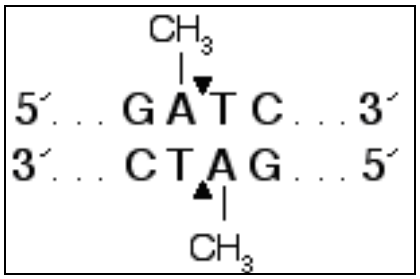

Figure 2.3: Restriction sites of *Dpn*I endonuclease

2.3.3 Competent cell preparation

For transformation of the mutagenesis products, BL21 electrocompetent cells were prepared in the laboratory. For the starting material commercial BL21 cells (New England Biolabs) were used. The preparation procedure was summarized in Figure 2. 4.

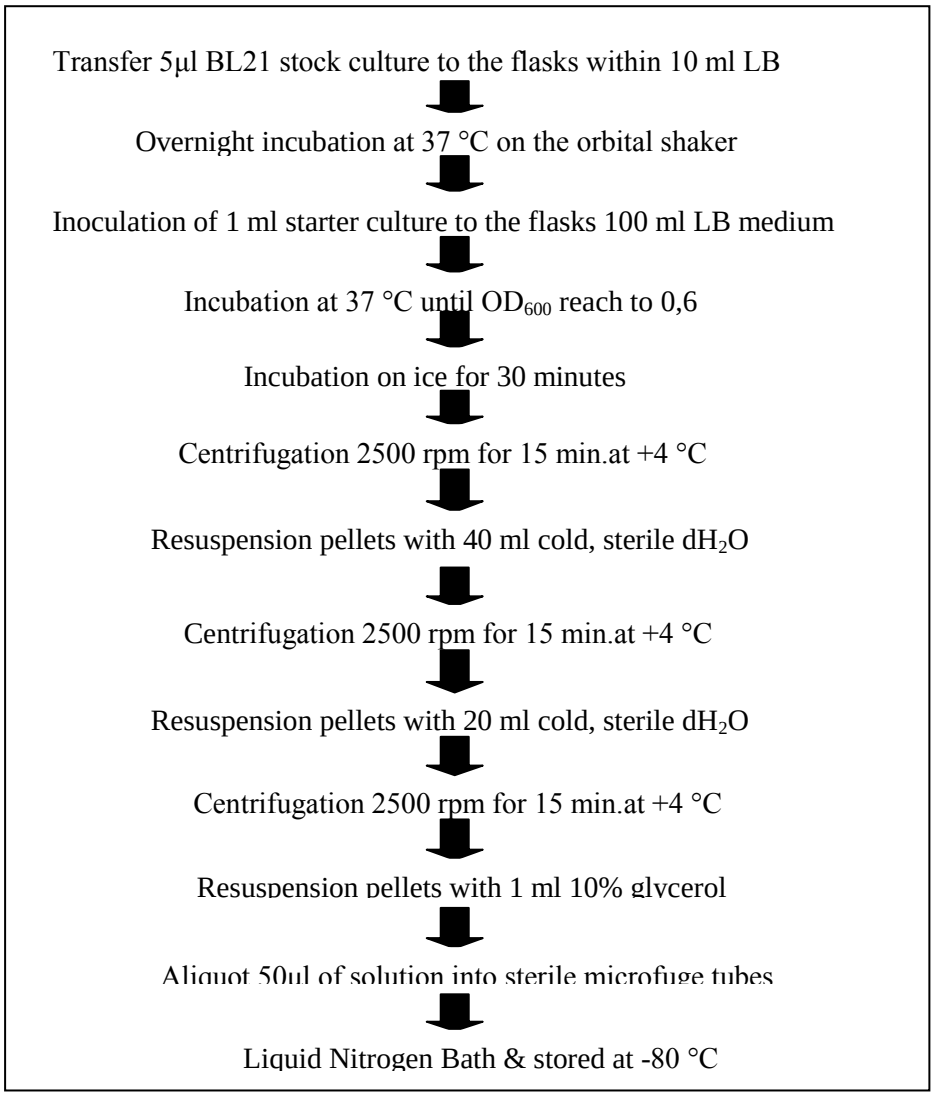

Figure 2.4: Scheme of the electrocompetent cell preparation prosedure

2.3.4 Transformation

After digestion with *Dpn*I, PCR products were transformed directly to previously prepared BL21 electrocompetent cells. For this purpose, 3 µl of PCR product was added into 50 µl BL21 electrocompetent cell, cells were then transferred to the electroporator cuvettes. After the electroporation, 250 µl S.O.C medium was immediately added to the cells, the solution was then transfered into tubes, and they were shaken and incubated for at least 1 hour at 37°C. Finally, after spreading 50 - 100 µl from each transformation on a prewarmed LB agar plate including (100 µg/ml) ampicillin, they were incubated overnight at 37°C.

2.3.5 Mutation confirmation

For the mutation confirmation, several colonies were picked randomly and cultured through overnight incubation at 37°C, then plasmid DNA was isolated as mentioned in section 2.3.1.

2.3.5.1 *Pst*I / *Sac*I restriction

For the mutation confirmation, firstly, isolated plasmid DNA was restricted by *Sac*I/*Pst*I restriction enzymes, in order to confirm *cm*FDH gene if inserted into *Sac*I/*Pst*I restriction sites of the PQE-2 vector. For this purpose, 10 µl of plasmid DNA was digested with *Sac*I: *Pst*I (1:1 units) for 2 hours at 37°C.

2.3.5.2 Sequence analysis

Sequence analysis was applied to the isolated plasmid DNA from the transformed colonies which were picked randomly in order to confirm the mutation. Sequence analysis composed of sequence PCR and sequence PCR product purification.

PCR conditions

For the sequence analysis, sequence PCR was applied by using BigDye Terminator V 3.1 Cycle Sequencing Kit (Applied Biosystems) with the primer 488F (5'-CTCTGTTGCTGAACACGTTC-3'). The ingredients and the conditions of the sequence PCR are shown in Table 2.4. and 2.5.

Table 2.7: Sequence PCR

Ingredient	Volume
5X Buffer	2 µl
RR-100	2 µl
Primer	1 µl
Template	1 µl
dH_2O	4 µl
Total	10 µl

Table 2. 8: Sequence PCR Conditions

Cycle Number	Degree	Time	Phase
1	96 °C	1 min	Initial Denaturation
30	96 °C	10 sec	Denaturation
	50 °C	5 sec	Annealing
	60 °C	4 min	Elongation
1	4°C	∞	Final Hold

PCR product purification

For the sequence analysis, PCR products were purified by the following procedure. Each 10 µl PCR product was transfered to 1,5 ml microfuge tubes, and 4 µl mixture of 3M Sodium Acetate:125mM EDTA and 50 µl 100% ethanol were added into the PCR products. After incubation at RT for 15 minutes, they were centrifugated at 13.000 x rpm for 15 minutes. The supernatant was discarded, then 70 µl 70% ethanol were added to the tubes and mixed well, and centrifugated at 13.000 x rpm for 15 minutes. After discarding the supernatant, the products were washed with 70 µl 70% ethanol, and centrifugated at 13.000 x rpm for 15 minutes. Finally, all supernatant was discarded and residual ethanol was removed, then the DNA was resuspended with 20 µl formamide and denaturated at 94 °C for a few minutes. The purified DNA was then analysed with cycle sequencer (ABI 3130 Avanti).

2.3.6 Library construction (stock culture preparation)

After site-saturation mutagenesis, transformed colonies, nearly 100 individual colonies from each library (D195, Y196, Q197, and D195/Y196/Q197) were picked

and transfered into 96-well microplates which included 100 µl Magic Media (MM) '*E.coli* Expression Medium' (Invitrogen) with ampicilin. MM is a commercial media which provides *E. coli* expression without IPTG induction. The master plates then were grown at 37°C for 18 hours, and were stored as 20% glycerol stocks frozen at -80°C for the further studies.

2.3.7 Screening with colorimetric assay

2.3.7.1 Lysate preparation

For screening the FDH activity, mutated individuals in the master plates of the libraries were transferred as triple replicates to the 96-deepwell microplates including 1,5 ml MM with ampicilin. Also each plate was included triple replication of the medium without any inoculation, PQE-2 vector without FDH insertion, and PQE-2 vector with wild type FDH as negative controls. After the deepwell plates were grown at 37°C for 22 hours, cells were harvested by centrifugation at 4000 x *g* for 20 minutes, and the medium was removed. The pellets then, were resuspended with 50 µl BugBuster (Novagen), and the plates were shaken at RT for 20 minutes. After incubation the plates were centrifugated at 4000 x *g* for 20 minutes, and then the lysates as supernatant were removed to new plates for the screening assays.

2.3.7.2 Activity measurement

For library screening, colorimetric assay was applied. Each 200 µl reaction contained 20 µl of lysate, 30 µl of dilution buffer (50 mM Tris-HCl [pH 8.0] and 0,13 % gelatin), and 150 µl reaction buffer (300 µM NADP, 300 µM formate, 300 µM NBT, and 30 µM PMS in dilution buffer). The reaction was started by adding lysate to the reaction solution, and then the kinetics were monitored at OD_{580} for the first 3 minutes. Finally, the reaction was observed in blue-purple color. The average of the triple replications of each mutant was analysed, and the candidate mutants with activity towards $NADP^+$ were determined. All these studies are summarized in Figure 2.5.

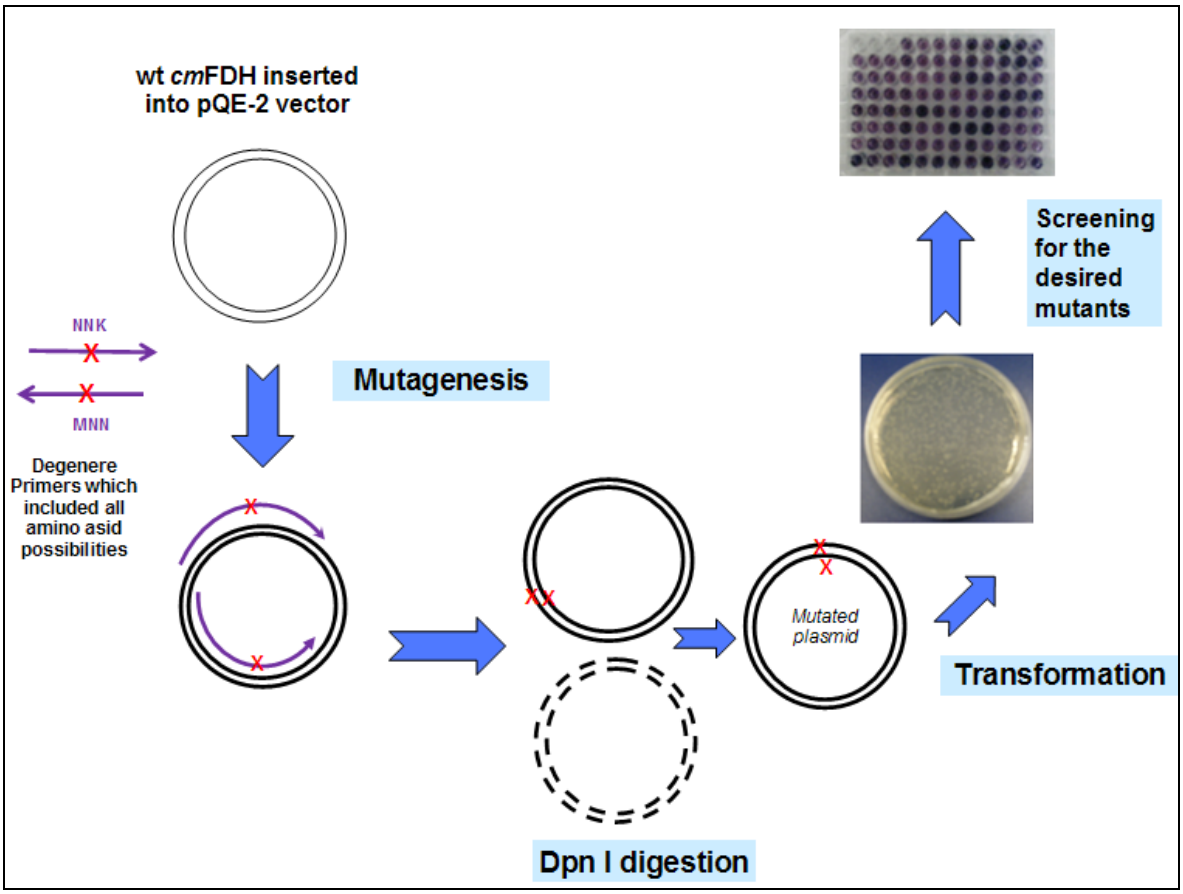

Figure 2.5: Summary of the experimental study of the site-saturation mutagenesis.

3. RESULTS AND DISCUSSION

3.1 Site-Saturation Mutagenesis Amplification

For the library construction of the determined NAD^+ binding residues, the diversity was set up with site-saturation mutagenesis which was mentioned in section 2.3.2. The amplified PCR products were then controlled in agarose gel electrophoresis, which is shown in Figure 3. 1.

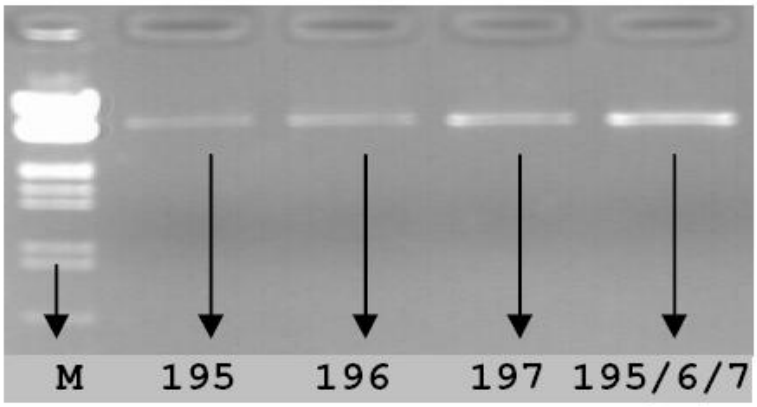

Figure 3.1: Agarose Gel Electrophoresis of determined residues in NAD^+ binding region. M; Marker 3(λ DNA/ *EcoR*I+ *Hind* III (Fermentas)); 21226 bp, 5184 bp, 4973 bp, 4268bp, 3530 bp, 2027 bp, 1904 bp, 1584 bp, 1375 bp, 947 bp, 831 bp, 564 bp., 195,196,197, and 195/6/7; are the PCR products of the mutated regions respectively. ≈ 6kb PCR products (≈ 5kb vector + ≈ 1kb the insert *cm*FDH) are shown respectively.

3.2 Transformation

The amplified templates with all amino acid possibilities at determined sites (D195, Y196, Q197, and D195/Y196/Q197) were transformed into BL21 electrocompetent cells. After transformation, 4 libraries D195, Y196, Q197, and D195/Y196/Q197 according to the mutated regions were constructed with many individuals (Figure 3.2, 3.3, 3.4, and 3.5).

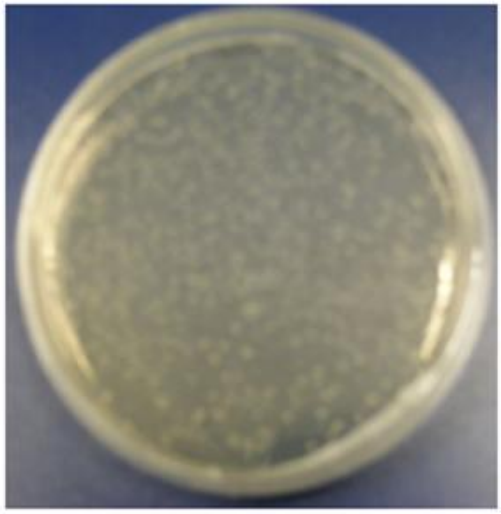

Figure 3.2: Library of D195

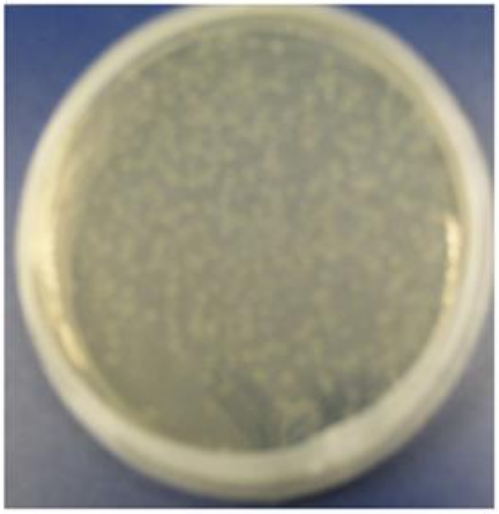

Figure 3.3: Library of Y196

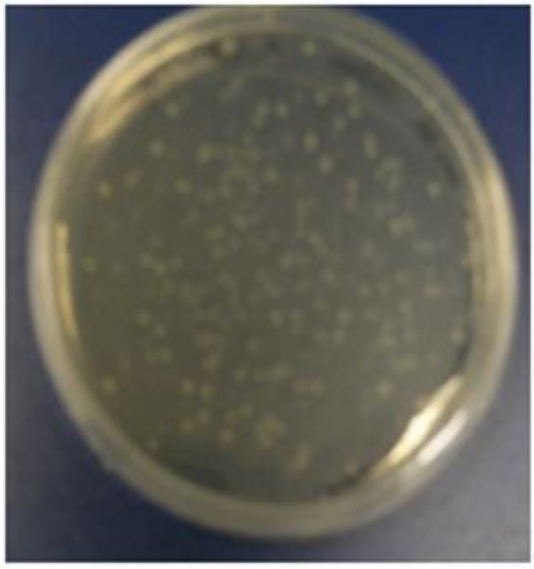

Figure 3.4: Library of Q197

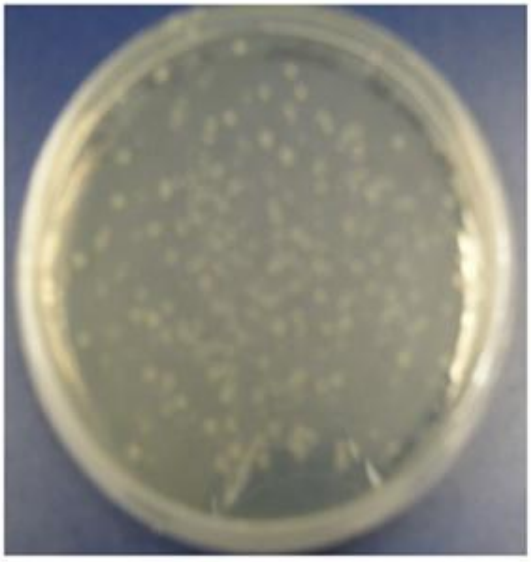

Figure 3.5: Library of D195/Y196/Q197

3.3 Mutation Confirmation

For the mutation confirmation, a few transformed mutants were picked, and analysed as mentioned in section 2. 6. *Sac*I/*Pst*I restriction was applied for the control of the insert *cm*FDH, and sequence analysis was applied for the control of the amino acid changes.

3.3.1 *Sac*I/*Pst*I restriction

Agarose gel electrophoresis of *Sac*I/*Pst*I restriction is shown in Figure 3.6.

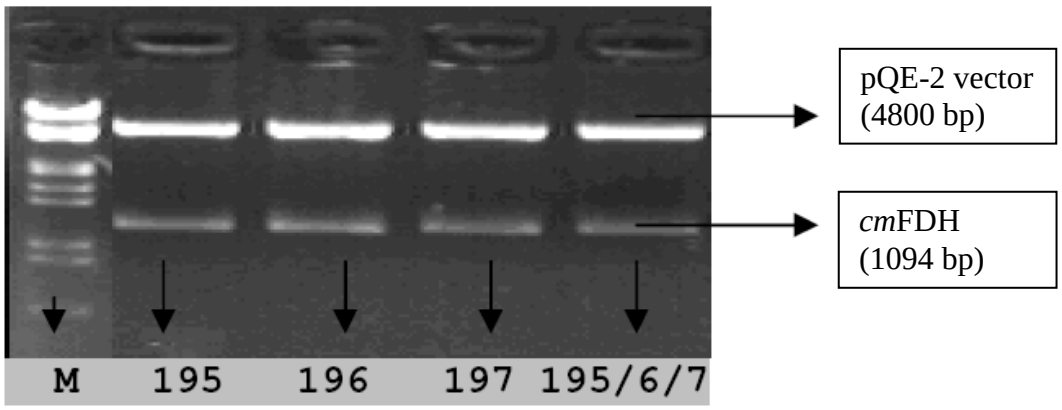

Figure 3.6: *Sac*I/*Pst*I restriction of the template DNA. M; Marker 3(λ DNA/ *EcoR*I+ *Hind* III (Fermentas)); 21226 bp, 5184 bp, 4973 bp, 4268bp, 3530 bp, 2027 bp, 1904 bp, 1584 bp, 1375 bp, 947 bp, 831 bp, 564 bp., 195,196,197, and 195/6/7; are the PCR products of the mutated regions respectively. ≈ 5kb vector, and ≈ 1kb the insert *cm*FDH are shown respectively.

3.3.2 Sequence analysis

The comparison of the sequences of the mutated individuals and the orgjinal *cm*FDH sequence is shown in Figure 3.7., the amino acid changes of the different residues are shown in Table 3.1.

```
                                                        195
  1          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACACTTATCAAGCT
  2          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACCGGTATCAAGCT
  3          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACCGTTATCAAGCT
  4          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACTTGTATCAAGCT
  5          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACATTTATCAAGCT
 fdh         GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATTATCAAGCT719
             ************************************************   *********

                                                      196
  1          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATGTGCAAGCT
  2          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATTTCAAGCT
  3          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATGCTCAAGCT
  4          GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATCATCAAGCT
 fdh         GTCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATTATCAAGCT719
             ************************************************   ******

                                                          197
  1          GTCTTGGAAAGATTACTCCGCATTTAATCCAAAAGAATTATTATACTACGATTATCCGGC
  2          GTCTTGGAAAGATTACTCCGCATTTAATCCAAAAGAATTATTATACTACGATTATGTGGC
 fdh         GTCTTGGAAAGATTACTCCGCATTTAATCCAAAAGAATTATTATACTACGATTATCAAGC718
             ************************************************************  **

                                                  195/196/197
  1          TCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGTGCGGCGGGCTT
  2          TCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGCTGGGGCGGCTT
  3          TCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACCGGCGTCAGGCTT
  4          TCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGGGCTTGCTGCTT
 fdh         TCTTGGAAAGATTACTCCCATTTAATCCAAAAGAATTATTATACTACGATTATCAAGCTT720
             ********************************************          ****
```

Figure 3.7: Sequences allignmets of the libraries (D195, Y196, Q197, and D195/Y196/Q197) respectively. The original residues are in blue color, the amino acid changes are shown in red.

Table 3.1: Amino acid changes in the determined residues (D195,Y196, Q197, and D195/Y196/Q197)

Amino acid residues	Amino acid Changes			
D195 (Aspartic Acid)	T	R	L	I
Y196 (Tyrosine)	V	F	A	H
Q197 (Glutamine)	P		V	
D195/Y196/Q197	V/R/R	A/G/A	R/R/Q	G/L/A

3.4 Screening with Colorimetric Assay

For screening the desired mutant in the libraries, activities of the individuals were detected indirectly by using colorimetric assay (NBT-PMS assay), as mentioned in section 2.8. Nearly 100 colonies of each library were screened at 580nm for 3 minutes. The increments of the optic density of the mutants and the negative controls are shown in Table 3.2. In the library of D195 ≈ 20 individuals, in the library of Y196 ≈ 30 individuals, in the library of Q197 ≈ 25 individuals, and in the library of (which all three residues were mutated) D195/Y196/Q197 ≈ 20 individuals are positive mutants, the rest of the individuals have shown no activity towards the $NADP^+$. The assay plates and the graphics of the mutants with higher activity compared with negative controls are shown in Figure 3.8- 3.15. The amino acid changes of some active mutants which were determined by cycle sequencing as mentioned in section 2.2.5, were shown in Table 3.3.

Table 3.2: Average OD of the mutants of D195, Y196, Q197, and D195/Y196/Q197 Libraries

Examples D195	Average OD	Examples Y196	Average OD	Examples Q197	Average OD	Examples 195/6/7	Average OD
Blank	0,000	Blank	0,000	Blank	-0,001	Blank	-0,001
PQE-2 without FDH	-0,001	PQE-2 without FDH	-0,001	PQE-2 without FDH	0,000	PQE-2 without FDH	0,000
PQE-2 with wild type FDH	0,001	PQE-2 with wild type FDH	0,000	PQE-2 with wild type FDH	-0,001	PQE-2 with wild type FDH	-0,024
M1	-0,034	M1	-0,033	M1	0,014	M1	-0,002
M2	0,008	M2	-0,014	M2	0,021	M2	0,013
M3	0,004	M3	-0,001	M3	0,035	M3	0,015
M4	-0,010	M4	-0,001	M4	-0,032	M4	0,017
M5	-0,019	M5	0,012	M5	0,007	M5	0,005
M6	0,005	M6	0,001	M6	0,012	M6	-0,004
M7	0,004	M7	-0,003	M7	-0,012	M7	-0,003
M8	-0,001	M8	-0,087	M8	0,028	M8	-0,004
M9	0,001	M9	-0,144	M9	0,016	M9	-0,029
M10	0,035	M10	-0,036	M10	0,013	M10	0,022
M11	0,006	M11	-0,033	M11	-0,112	M11	-0,018
M12	0,006	M12	-0,018	M12	0,004	M12	-0,011
M13	0,005	M13	-0,051	M13	0,016	M13	-0,063
M14	-0,008	M14	-0,007	M14	-0,075	M14	0,010
M15	0,008	M15	-0,023	M15	-0,049	M15	0,007
M16	0,005	M16	-0,002	M16	0,009	M16	0,008
M17	0,003	M17	-0,005	M17	0,032	M17	-0,005
M18	0,005	M18	0,007	M18	0,003	M18	-0,004
M19	0,007	M19	-0,004	M19	0,069	M19	0,026
M20	0,005	M20	0,021	M20	0,013	M20	0,001
M21	0,000	M21	-0,003	M21	0,019	M21	-0,009
M22	-0,059	M22	-0,117	M22	-0,014	M22	-0,005
M23	-0,011	M23	-0,142	M23	-0,035	M23	0,055
M24	-0,001	M24	0,007	M24	0,008	M24	0,052
M25	-0,082	M25	-0,002	M25	0,002	M25	0,001
M26	0,004	M26	0,003	M26	0,000	M26	-0,024
M27	0,007	M27	0,019	M27	-0,078	M27	0,006
M28	0,008	M28	0,002	M28	0,056	M28	0,059
M29	0,091	M29	-0,010	M29	-0,052	M29	0,021
M30	0,002	M30	-0,002	M30	-0,108	M30	0,008
M31	-0,050	M31	0,002	M31	0,012	M31	0,000
M32	0,015	M32	0,008	M32	0,010	M32	0,013
M33	0,015	M33	0,000	M33	0,013	M33	-0,050
M34	0,014	M34	-0,038	M34	0,016	M34	0,008
M35	-0,028	M35	0,008	M35	-0,069	M35	0,005
M36	-0,020	M36	-0,031	M36	-0,101	M36	0,004
M37	0,000	M37	-0,016	M37	0,031	M37	0,027
M38	0,000	M38	-0,006	M38	0,071	M38	0,022
M39	-0,007	M39	0,009	M39	0,011	M39	-0,027
M40	-0,019	M40	-0,035	M40	-0,060	M40	-0,004
M41	0,001	M41	0,017	M41	0,013	M41	0,046
M42	0,019	M42	0,013	M42	0,020	M42	0,019
M43	0,010	M43	0,011	M43	0,034	M43	-0,014

Table 3.2 (continued): Average OD of the mutants of D195, Y196, Q197, and D195/Y196/Q197 Libraries

M44	0,004	M44	0,012	M44	0,007	M44	-0,044
M45	0,010	M45	0,008	M45	0,015	M45	0,055
M46	0,011	M46	0,020	M46	0,035	M46	0,003
M47	0,025	M47	0,012	M47	0,022	M47	-0,080
M48	0,038	M48	0,011	M48	-0,003	M48	0,005
M49	0,020	M49	0,017	M49	0,014	M49	-0,034
M50	0,020	M50	0,025	M50	0,026	M50	0,026
M51	0,036	M51	-0,027	M51	0,002	M51	-0,007
M52	0,028	M52	0,020	M52	-0,008	M52	0,005
M53	0,064	M53	0,029	M53	-0,024	M53	0,039
M54	0,014	M54	-0,028	M54	-0,043	M54	0,019
M55	0,047	M55	0,013	M55	-0,018	M55	-0,065
M56	0,044	M56	0,011	M56	-0,005	M56	0,001
M57	0,045	M57	0,028	M57	0,010	M57	0,002
M58	0,031	M58	0,008	M58	-0,051	M58	0,016
M59	0,013	M59	0,039	M59	0,020	M59	-0,030
M60	-0,022	M60	-0,046	M60	-0,012	M60	-0,014
M61	0,016	M61	0,004	M61	0,012	M61	0,028
M62	0,008	M62	0,039	M62	-0,025	M62	0,011
M63	0,003	M63	0,036	M63	0,004	M63	0,017
M64	-0,038	M64	-0,033	M64	-0,023	M64	-0,035
M65	0,012	M65	0,028	M65	0,002	M65	-0,054
M66	0,004	M66	0,043	M66	0,006	M66	0,016
M67	0,005	M67	0,044	M67	0,008	M67	0,037
M68	0,003	M68	0,026	M68	0,004	M68	-0,009
M69	0,025	M69	0,031	M69	0,011	M69	-0,047
M70	0,004	M70	0,041	M70	0,029	M70	-0,011
M71	0,001	M71	0,039	M71	0,070	M71	-0,016
M72	0,006	M72	0,011	M72	0,009	M72	-0,012
M73	0,002	M73	0,022	M73	0,013	M73	0,051
M74	0,004	M74	0,032	M74	0,062	M74	0,018
M75	0,016	M75	0,029	M75	0,049	M75	0,013
M76	0,012	M76	-0,032	M76	0,010	M76	0,017
M77	0,002	M77	-0,025	M77	-0,031	M77	0,014
M78	-0,024	M78	0,040	M78	0,020	M78	0,062
M79	-0,011	M79	0,053	M79	-0,047	M79	0,069
M80	0,035	M80	0,026	M80	-0,075	M80	0,031
M81	0,021	M81	0,020	M81	0,047	M81	0,011
M82	0,013	M82	0,027	M82	0,019	M82	0,010
M83	0,011	M83	0,039	M83	0,036	M83	0,010
M84	0,022	M84	0,004	M84	0,012	M84	-0,043
M85	0,011	M85	0,030	M85	0,035	M85	-0,077
M86	0,010	M86	0,044	M86	0,029	M86	0,006
M87	-0,046	M87	0,053	M87	0,014	M87	0,020

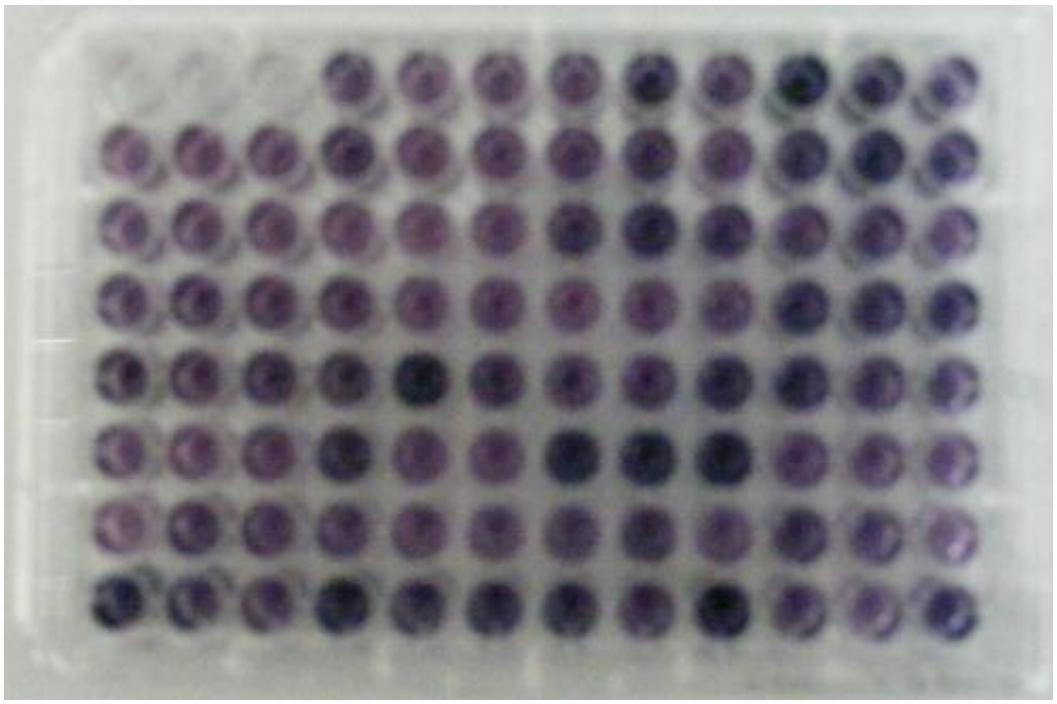

Figure 3.8: NBT-PMS assay reaction of the D195. The first 9 wells are the negative controls (the first triplet is blank, the second triplet is empty PQE-2, the third triplet is PQE-2 with *cm*FDH), and the rest of them are mutants.

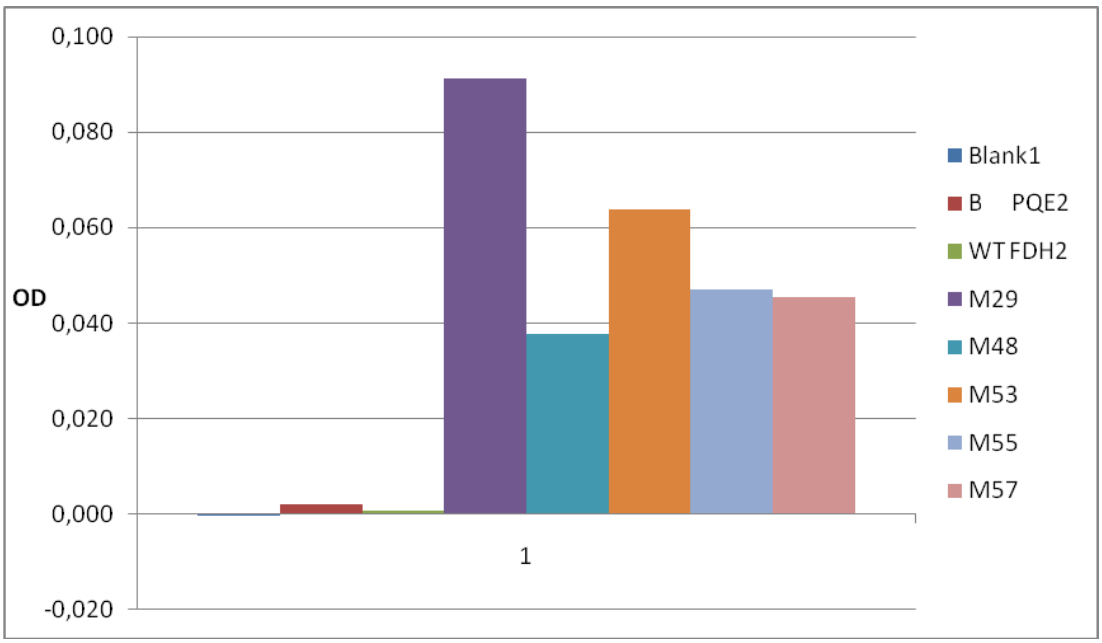

Figure 3.9: Graphic of the D195 screening assay. The negative controls and the positive mutants are compared.

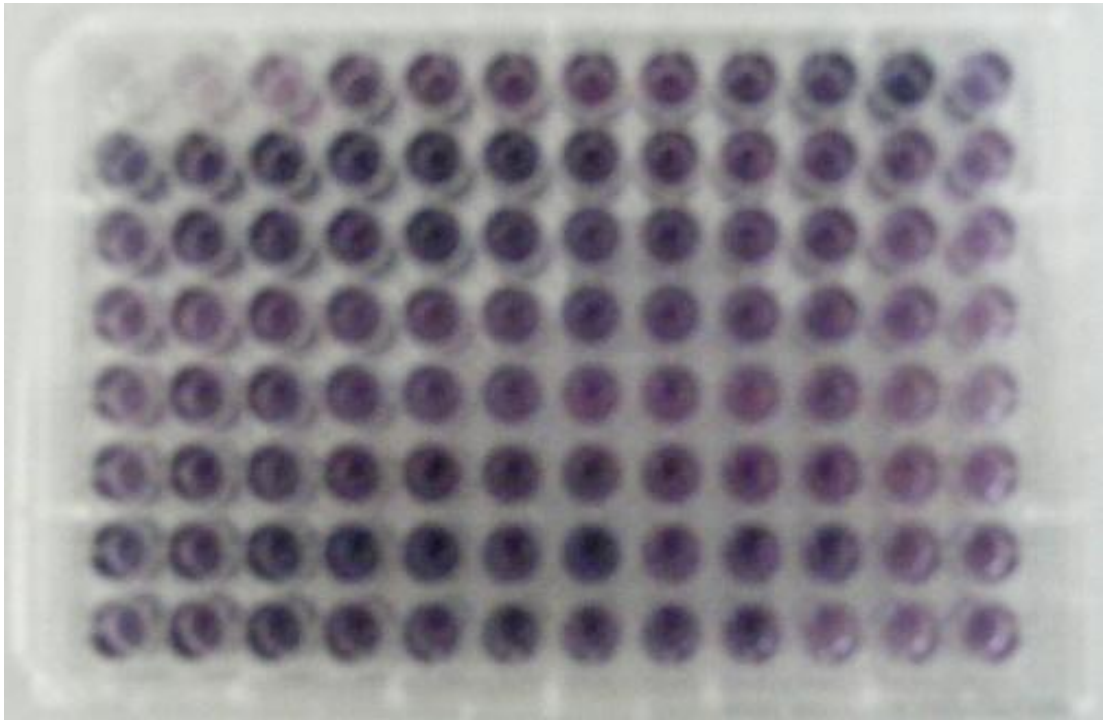

Figure 3.10: NBT-PMS assay reaction of the Y196. The first 9 wells are the negative controls (the first triplet is blank, the second triplet is empty PQE-2, the third triplet is PQE-2 with *cm*FDH), and the rest of them are mutants.

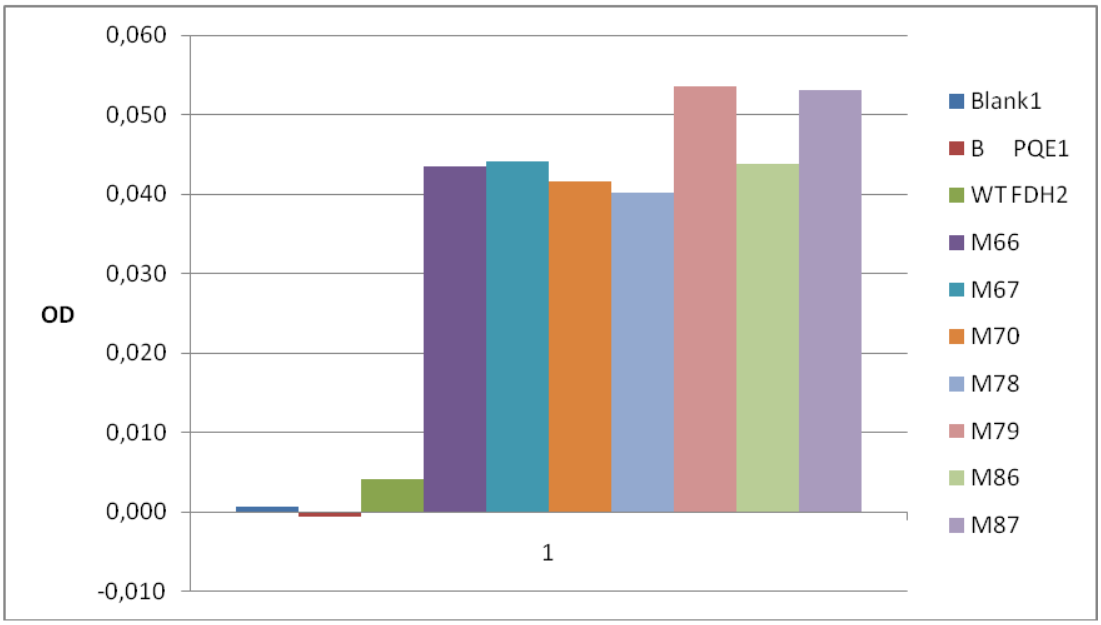

Figure 3.11: Graphic of the Y196 screening assay. The negative controls and the positive mutants are compared.

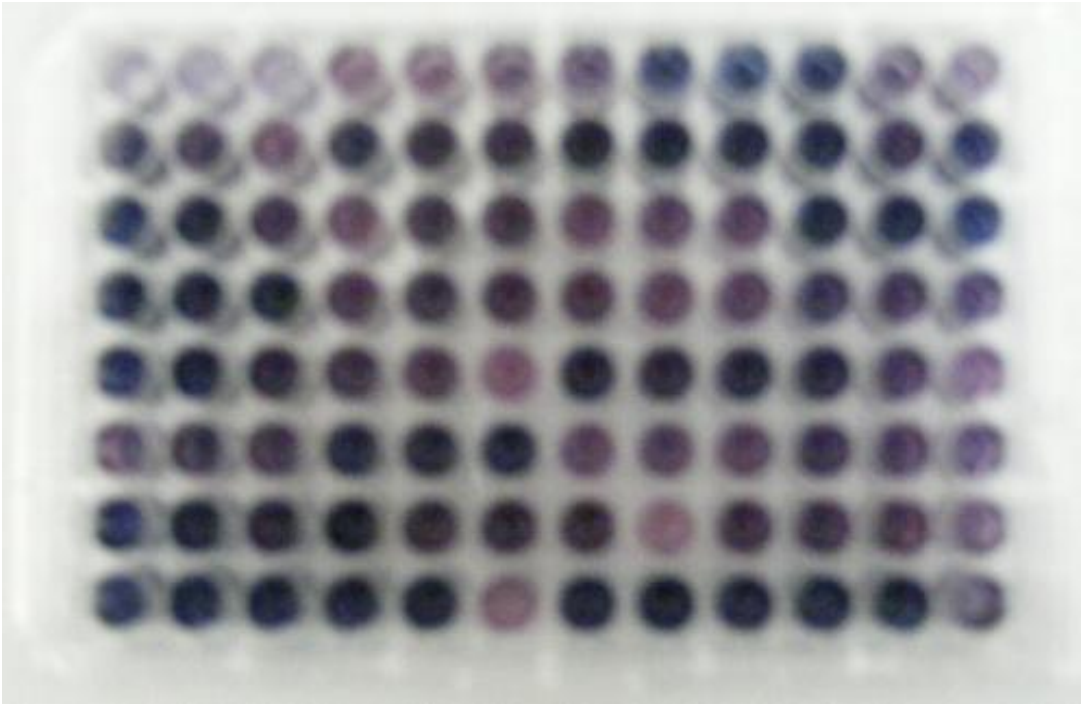

Figure 3.12: NBT-PMS assay reaction of the Q197. The first 9 wells are the negative controls (the first triplet is blank, the second triplet is empty PQE-2, the third triplet is PQE-2 with *cm*FDH), and the rest of them are mutants.

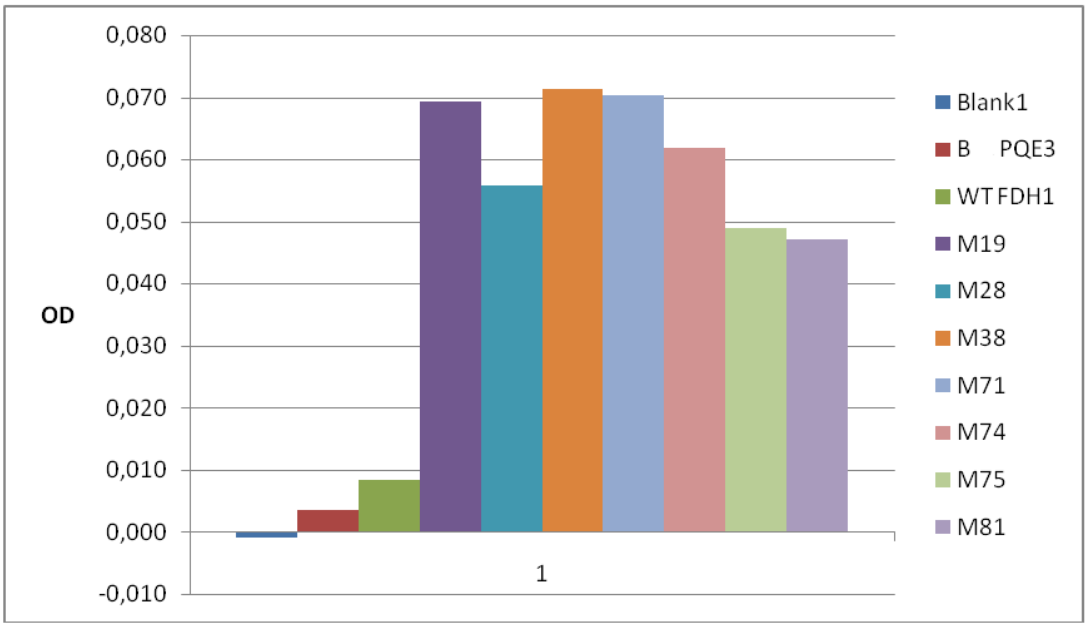

Figure 3.13: Graphic of the Q197 screening assay. The negative controls and the positive mutants are compared.

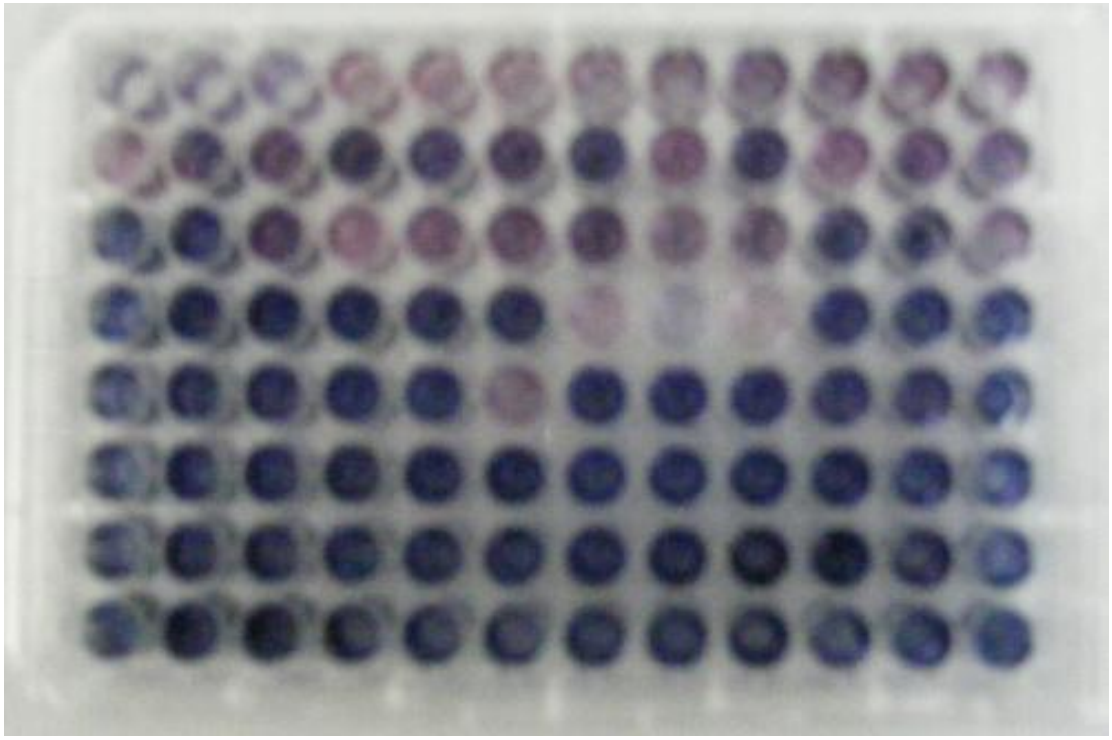

Figure 3.14: NBT-PMS assay reaction of the D195/Y196/Q197. The first 9 wells are the negative controls (the first triplet is blank, the second triplet is empty PQE-2, the third triplet is PQE-2 with *cm*FDH), and the rest of them are mutants.

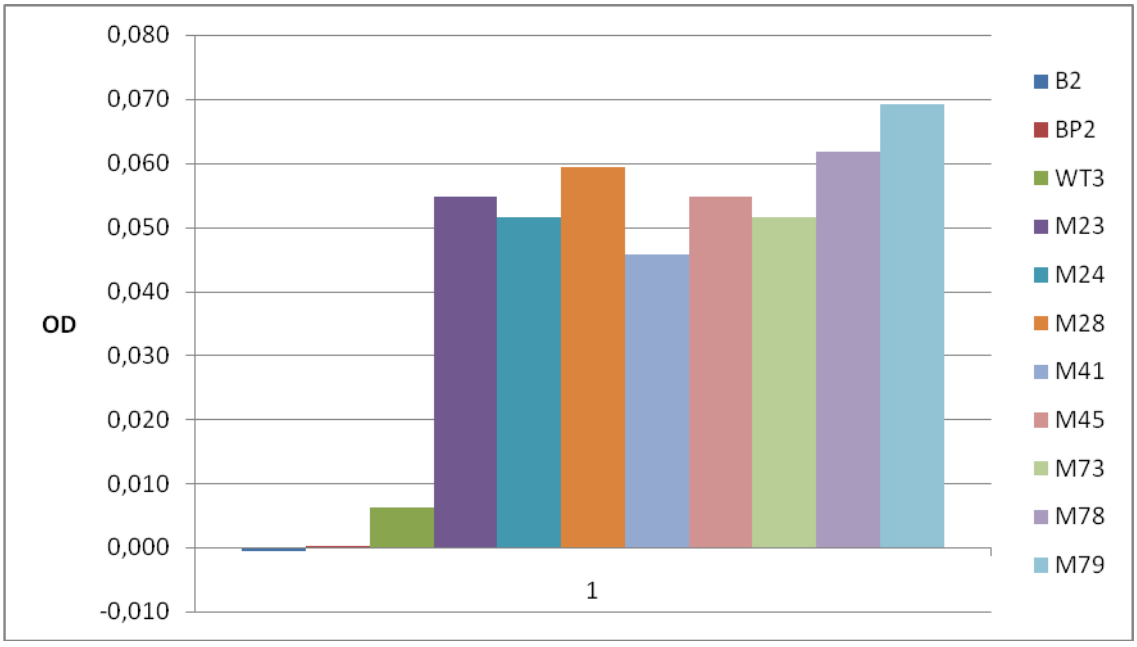

Figure 3.15: Graphic of the D195/Y196/Q197 screening assay. The negative controls and the positive mutants are compared.

Table 3.3: Amino acid changes of the active mutants in the determined residues (D195,Y196, Q197, and D195/Y196/Q197).

Amino acid residues	Names of mutants	Amino acid changes
D195 (Aspartic Acid)	195/M29	D195S
	195/M53	D195F
	195/M55	D195G
	195/M56	D195C
Y196 (Tyrosine)	196/M66	Y196W
	196/M70	Y196V
	196/M86	Y196H
Q197 (Glutamine)	197/M19	Q197V
	197/M28	Q197M
	197/M38	Q197P
	197/M75	Q197A
D195/Y196/Q197	567/M41	D195Y/Y196C/Q197R

4. CONCLUSION

NAD^+-dependent formate dehydrogenase (EC 1.2.1.2, FDH) is the last enzyme in the metabolism of methanol in methylotrophs and catalyzes the oxidation of the formate anion into carbon dioxide concomitant with the reduction of NAD^+ to NADH. One crucial role of this enzyme in industrial redox chemistry is to regenerate NADH from NAD^+- and formate. The amino acid sequences of FDHs from aerobic to anaerobic organisms have been reported and there are crystal structures of FDHs from the bacterium *Pseudomonas sp* 101 (*ps. sp*101) and *Candida boidinii* (*cb*). An investigation of FDHs led to the selection of C*andida* species FDH as the best candidate for the NADH regeneration system, because it is stable and it has relatively good activity. FDH from *Candida methylica* (*cm*) was cloned and overproduced at the University of Bristol and purification processes have been improved at Department of Molecular Biology and Genetics of Istanbul Technical University (ITU) giving a much better yield. The one disadvantage of FDH is that it uses only NAD^+ as a coenzyme. It would be also desirable to regenerate NADPH by using $NADP^+$ as a coenzyme. Many attempts using a rational design approach have been made to change the coenzyme specificity of FDH from NAD^+ to $NADP^+$ but all these mutants bind $NADP^+$ very weakly and still show activity with NAD^+.

In this project, we used site saturation mutagenesis which is a technique using a directed evolution approach to redesign proteins to improve the K_M of *cm*FDH for $NADP^+$. Firstly, in the coenzyme binding domain the amino acid residues which are responsible for the coenzyme specificity were determined by using Insight II (Accelrys) program on a homology model of *cm*FDH based on *ps. sp*.101 and *cb*. After the application of site saturation mutagenesis with degenerate primers to the determined residues D195, Y196, Q197, mutant libraries were constructed. The efficient mutants have been screened by using suitable screening assays in 96-well microplates.

According to the relative results of the colorimetric assay we obtained a set of positive mutants that show clearly improved $NADP^+$ specificity compared to the

wild-type. Following the production of both wild type and mutant FDHs, kinetic studies of positive mutants will be performed. Also further site-saturation mutagenesis can be applied in order to construct further generations of the efficient mutants which can be used as templates to improve the properties of the $NADP^+$-specific *cm*FDH.

In conclusion, site-saturation mutagenesis of crucial residues in NAD(P) specificity, D195, Y196 and Q197 from *cm*FDH coupled with screening by colorimetric assay has resulted in promising candidate mutants with high $NADP^+$ activity. 27 candidates have shown activity towards $NADP^+$ from nearly 400 screened colonies. These results reveal that, these residues are important in controlling the cofactor specificity of formate dehydrogenases, and promising for developing $NADP^+$-dependent *cm*FDH enzyme.

Further kinetic assays will be applied in order to determine the most active and effective candidate for the NADPH regeneration, also further generations can be constructed using the active mutant as a template to improve the efficiency of the $NADP^+$-dependent FDH for industrial usage.

.

REFERENCES

[1] Andreadeli, A., Platis, D., Tishkov, V. I., Popov, V. O., Labrou, N. E., 2008. Structure-guided alteration of coenzyme specificity of formate dehydrogenase by saturation mutagenesis to enable efficient utilization of $NADP^+$, *FEBS Journal*

[2] Antikainen, N. M., Martin, S. F., 2005: Altering protein specificity: techniques and applications, *Bioorganic & Medicinal Chemistry*, **13,** 2701–2716

[3] Arnold, F. H., Georgiou, G., 2003. DirectedEnzyme Evolution: Screening and Selection Methods, Humana Press Inc., Totowa, NJ

[4] Arnold, F.H., 2007. Directed enzyme evolution, http://www.che.caltech.edu/groups/fha/directed_evolution.html

[5] Cedrone, F., Ménez, A., Quéméneur, E., 2000. Tailoring new enzyme functions by rational redesign, *Current Opinion in Structural Biology*, **10,** 405–410

[6] Chen, R., 2001. Enzyme engineering: rational redesign versus directed evolution, *Trends in Biotechnology*, **19,** 13-14

[7] Chen,R., Greer, A., Dean, A. M., 1996. Redesigning secondary structure to invert coenzyme specificity in isopropylmalate dehydrogenase, *Proc. Natl. Acad.Sci.* U. S. A., **93,** 12171–12176

[8] Chen, R., 1999. A general strategy for enzyme engineering, *Trends Biotechnol*., **17,** 344–345

[9] Chica, R. A., Doucet, N., Pelletier, J. N., 2005. Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design, *Current Opinion in Biotechnology*, **16,** 378–384

[10] Crameri, A., Raillard, S. A., Bermudez, E., Stemmer, W. P., 1998. DNA shuffling of a family of genes from diverse species accelerates directed evolution, *Nature*, **391,** 288-291

[11] Debnam, P. M., Shearer, G., 1997. Colorimetric Assay for Substrates of $NADP^+$ -Dependent Dehydrogenases Based on Reduction of a Tetrazolium Dye to Its Soluble Formazan, *Analytical Biochemistry*, **250,** 253-255

[12] DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z., Tomlinson, G., Huang, H., Tan, X., Bibbs, L., Chen, P., 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM), *J. Am. Chem. Soc*., **125,** 11476-11477

[13] Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., Hallewell, R. A., 1992. Faster superoxide dismutase mutants designed by enhancing electrostatic guidance, Nature, **358,** 347–351

[14] Hibbert, E. G., Baganz, F., Hailes, H. C., Ward, J. M., Lye, G. J., Woodley, J. M., Dalby, P.A., Directed evolution of biocatalytic processes, 2005. *Biomolecular Engineering*, **22,** 11–19

[**15] Hurley, J. H., Chen, R., Dean, A. M.,** 1996. Determinants of cofactor specificity in isocitrate dehyrdrogenase: structure of an engineered $NADP^+ \rightarrow NAD^+$ specificity-reversal mutant, *Biochemistry*, **35,** 5670–5678

[16] Johannes, T. W., Zhao, H., 2006. Directed evolution of enzymes and biosynthetic pathways, *Current Opinion in Microbiology*, **9,** 261–267

[17] Johannes, T. W., Woodyer, R. D., Zhao, H., 2007. Efficient Regeneration of NADPH Using an Engineered Phosphite Dehydrogenase, *Biotechnol. Bioeng*., **96,** 18–26

[18] Karaguler, N. G., Sessions, R.B., Clarke, A. R., Holbrook, J. J., 2001. A single mutation in the NAD-specific formate dehydrogenase from *Candida metylica* allows the enzyme to use NADP, *Biotecnology Letters*, **23,** 283-287

[19] Kelly, G. C., Chen, R. R., 2007. A Window into Biocatalysis and Biotransformations, *Biotechnol. Prog*., **23,** 52-54

[20] Lamzin, V.S., Dauter, Z., Popov, V.O., Harutyunyan, E.H., Wilson, K.S., 1994. High resolution structures of holo and apo formate dehydrogenase, *J. Mol.Biol*., **236,** 759–785

[21] Leisola, M., Turunen, O., 2007. Protein engineering: opportunities and challenges, *Appl Microbiol Biotechnol.*

[22] Mayer, K. M., Arnold, F. H., (2002), A Colorimetric Assay to Quantify Dehydrogenase Activity in Crude Cell Lysates, *The Society for Biomolecular Screening*, 7: 135-140

[23] McLachlan, M. J., Johannes, T. W., Zhao, H., 2008. Further improvement of phosphite dehydrogenase thermostability by saturation mutagenesis, *Biotechnol. Bioeng*., **99,** 268-274

[24] Mills, D. R., Peterson, R. L., Spiegelman, S., 1967. An extracellular Darwinian experiment with a self-duplicating nucleic acid molecule, *Proc. Natl. Acad. Sci. U.S.A.*, **58,** 217-224

[25] Miyazaki, K., Arnold, F. H., 1999. Exploring nonnatural evolutionary pathways by saturation muatgenesis: rapid improvement of protein function, J. Mol. Evol., **49,** 716-720

[26] Parikh, M. R., Matsumura, I., 2005. Site-saturation mutagenesis is more efficient than DNA shuffling for the directed evolution of β-Fucosidase from β- Galactosidase, *J. Mol. Biol*., **352,** 621-628

[27] Pitela, S. B., Chob, C. M., Chen, W., Zhao, H., 2007. Chapter 3. Directed Evolution Tools in Bioproduct andBioprocess Development, *Bioprocessing for Value-Added Products from Renewable Resources*, 49-72

[28] Popov, V. 0., Lamzin, V. S., 1994. NAD+-dependent formate dehydrogenase, *Biochem. J.*, **301,** 625-643

[29] Que´me´ neur, E., Moutiez, M., Charbonnier, J. B., Me´ nez, A., 1998. Engineering cyclophilin into a proline-specific endopeptidase, *Nature*, **391,** 301–304

[30] Rubingh, D. N., 1997. Protein engineering from a bioindustrial point of view, *Current Opinion in Biotechnology*, **8,** 417-422

[31] Schirwitz, K., Schmidt, A., Lamzin, V. S., 2007. High-resolution structures of formate dehydrogenase from Candida boidinii, *Protein Sci.* **16,** 1146-1156

[32] Seelbach, K., Riebel, B., Hummel, W., Kula, M. R., Tishkov, V., Egorov, A. M.,Wandrey, C., Kragl, U., 1996. A novel, efficient regenerating method of NADPH using a new formate dehydrogenase, *Tetrahedron Lett*., **37,** 1377–1380

[33] Serov, A. E., Popova, A. S., Fedorchuk, V. V., Tishkov, V. I., 2002. Engineering of coenzyme specificity of formate dehydrogenase from *Saccharomyces cerevisiae*, *Biochem. J*., **367,** 841–847

[35] Spadiut, O., Pisanelli, I., Maischberger, T., Peterbauer, C., Gorton, L., Chaiyen, P., Haltrich, D., 2008. Engineering of pyranose 2-oxidase: Improvement for biofuel cell and food applications through semi-rational protein design, *Journal of Biotechnology*, **139,** 250–257

[36] Stemmer, W. P., 1994. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution, *Proc. Natl. Acad. Sci. U.S.A.*, **91,** 10747-10751

[37] Tishkov, V. I., Popov, V. O., 2004. Catalytic Mechanism and Application of Formate Dehydrogenase, *Biochemistry*, **69,** 1252-1267

[38] Tishkov, V. I., Popov, V. O., 2006. Protein engineering of formate dehydrogenase, *Biomolecular Engineering*, **23,** 89–110

[39] Wahler, D., Reymond, J. L., 2001. Novel methods for biocatalyst screening, *Current Opinion in Chemical Biology*, **5,** 152-158

[40] Williams, G. J., Nelson, A. S., Berry, A., 2004. Directed evolution of enzymes for biocatalysis and the life sciences, *Cell. Mol. Life Sci*., **61,** 3034–3046

[41] Wong, T. S., Tee, K. L., Hauer, B., Schwaneberg, U., 2004. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution, *Nucleic Acids Res*,

[42] Yoshikuni, Y., Ferrin, T. E., Keasling, J. D., 2006. Designed divergent evolution of enzyme function, *Nature*, **440,** 1078–1082

[43] Yuan, L., Kurek, I., English, J., Keenan, R., Laboratory-Directed Protein Evolution, 2005. *Microbiology and Molecular Biology Reviews*, **69,** 373–392

[44] Zhao, H., Zha, W., 2003. ‘‘Evolutionary Methods for Protein Engineering.” In Enzyme Functionality: Design, Engineering and Screening, , Marcel Dekker, Inc., New York, NY, pp. 353-373

[45] Rozzell, J. D., Hua, L., Mayhew, M., Novick, S., 2004. Mutants of enzymes and methods for their use, *US Patent Application Publication*, No: US2004/0115691

[46] Ordu, E.B., Karagüler, N.G., 2007. Improving the purification of NAD^+dependent formate dehydrogenase from *Candida methylica*, *Prep. Biochem. and Biotech.*, **37,** 333-341

CURRICULUM VITAE

Candidate's full name: Gulsah PAR

Place and date of birth: Bornova/ IZMIR – 21/08/1981

Permanent Address: Tarabya/ ISTANBUL

Universities and Colleges attended: B.Sc. : Biology, Egean University (1999-2004)

Publications:

- **Par, G.,** Yelboga, E., Ordu, E.B., Sessions, R.B., Clarke, A. R, Karaguler, N. G., Site-Saturation Mutagenesis of Crucial Residues in NAD /NADP specificity of Formate Dehydrogenase from *Candida methylica*, 8th International ProStab Conference, poster presentation, in 2009, in Graz (AUSTRIA)

- Karaguler, N. G., Ordu, E., Yelboga, E., **Par, G.,** Binay, B.,. Sessions, R. B., Clarke, A. R., Nanobiyokatalizör Geliştirilmesinde Protein Mühendisliği Uygulamaları, Istanbul Technical Universty, NanoScience and Nanotechnology Workshop, poster presentation, in 2009, in Istanbul (TURKEY)